THE UNIVERSITY OF ALBERTA

CLONING OF PMA-INDUCED cDNA SEQUENCES FROM EL4 MOUSE LYMPHOMA CELLS

by

JOHN FRANCIS ELLIOTT

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

EDMONTON, ALBERTA

SPRING 1987

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-37614-7

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR:	John Francis Elliott
TITLE OF THESIS:	Cloning of PMA-Induced cDNA Sequences from EL4 Mouse Lymphoma Cells
DEGREE FOR WHICH THESIS WAS PRESENTED:	Doctor of Philosophy
YEAR THIS DEGREE GRANTED:	Spring 1987

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) .....

PERMANENT ADDRESS:

8609 - 111 St

Edmonton, Alta.

T6G 1H8

DATED Dec. 18 1986

'Before long, a post-M.D. period of training in internal medicine and molecular genetics might serve to prepare a young graduate for almost any discipline, in the kind of medicine that lies somewhere ahead.'

Lewis Thomas, from the introduction to the *Cecil Textbook of Medicine* , 16th Edition, 1982.

'Make sure you have the proper scientific attitude, be strictly honest, search for the truth and try not to make yourself a reputation, but to contribute to knowledge. The expectation must be that you will make an important contribution, and develop a reputation accordingly.'

Raymond Lemieux, interviewed in the *Alberta Heritage Foundation for Medical Research Newsletter,* Summer 1986.

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled CLONING OF PMA-INDUCED cDNA SEQUENCES FROM EL4 MOUSE LYMPHOMA CELLS submitted by John F. Elliott in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry.

Verner Paetkau
Supervisor

.....

D G Scraba

Wm Paranchych

Bhagirath Singh

J Walter Gallop
External Examiner

Date Dec 8/86

To my friend Anne.

ABSTRACT

On exposure to phorbol-12-myristate-13-acetate (PMA) the mouse T lymphoma cell line EL4 begins to secrete a number of lymphokines which act on T and B lymphocytes. The first part of this work demonstrates that in EL4 cells the PMA-mediated induction of lymphokine mRNA's coding for interleukin 2 (IL2) and colony stimulating factor for granulocyte-macrophages (GM-CSF) is selectively blocked in the presence of the immunosupressive agent cyclosporin A (CsA). In the second part of this study, a number of cDNA clones containing sequences which are induced by PMA in EL4 cells were isolated by using the method of subtractive hybridization, both to clone a subtracted cDNA library (enriched for PMA-induced sequences) and to generate subtracted cDNA probes. Of the 136 individual clones isolated, about 75% contained essentially the same sequence, which is derived from the *env* and 3'LTR region of mouse mammary tumor virus (MMTV). This PMA-inducible MMTV transcript is blocked in the presence of CsA, and cap labelling experiments demonstrate that the transcript initiates within the *env* region of the provirus. Several other PMA-inducible sequences occurred at a much lower frequency (1-5 clones in the group of 136 clones). Sequence analysis allowed the identification of three of these, including mouse IL2, mouse carbonic anhydrase II, and the precursor protein for mouse chondroitin sulfate. The expression of the latter two transcripts was unaffected by CsA. Finally, a PMA-inducible sequence was characterized which could not be found in the literature. The mRNA has a characteristic 8 base sequence in the 3' untranslated region which suggests that it may code for a new lymphokine or growth factor (the secreted protein is about 115 amino acids in length). This transcript is 'superinduced' in EL4 in the presence of PMA and CsA, and is also induced in another helper T-cell line following mitogen activation. The cDNA was placed in an SV40 derived expression vector and transfected into COS monkey cells, and the resulting product was tested for biological activity in a number of *in vitro* immunological assays.

ACKNOWLEDGEMENTS

I would like to acknowledge the support of my supervisor, Dr. Verner Paetkau, and my supervisory committee including Drs. D. Scraba, W. Paranchych and B. Singh. Several other members of the Biochemistry Department, including Drs. R. Morgan, G. McFadden, M. James and L. Smillie have been especially friendly and helpful, and I am grateful for their time and interest.

A number of individuals from other universities have been instrumental in teaching me many of the techniques which are described in Chapter II, and I am particularly indebted to them. They are: Y. Chien, M.M. Davis, D. Denny, N. Gascoigne, S. Gillam, J. Hagmann, G. O'Neill, and G. Pielak. I would also like to acknowledge the gift of oligonucleotide probes from P. Barr, the SV40 expression plasmid from J. Miller, and the Cyclosporin A from F. Pasutto.

I am grateful for expert technical assistance from C. Gibbs and in particular from R. Garner, who prepared some of the RNA and did some of the DNA sequencing described in this thesis. J. Hooton and C. Havele generously offered advice about the tissue culture and also supplied many of the cell lines. G. Ritzel did most of the oocyte injections described in chapter III, and patiently taught me the technique. D. Harnish, J. Shaw, and C. Bleackley collaborated in the early stages of this work, and in particular D. Harnish introduced me to the techniques of molecular biology and originally suggested the research project described in this thesis. D. Talbot and B. Pohajdak contributed a great deal of time and effort to the latter half of the project, and I gratefully acknowledge their help. I am particularily thankful to B. Pohajdak for his continuing good humor and friendship through some rather difficult times. My fellow graduate student C. Macaulay suggested the cap labelling experiments described in chapter VIII, and was a helpful sounding board on a number of occasions. I also

gratefully acknowledge the cooperation of C. Eaves, D. Kunimoto, C. Havele, J. Plate, A. Greenberg, D. Green, and L. Guilbert in testing the COS monkey cell supernatants for biological activity.

The Alberta Heritage Foundation for Medical Research has given me very generous financial support through the latter half of my medical training and during my graduate studies; without this support I would not have been able to pursue a career in medical research.

Finally I would like to acknowledge the continuous interest and encouragement of my wife Anne Macdonald, as well as the moral support of the Elliott, Macdonald and Asch families.

TABLE OF CONTENTS

	Page
CHAPTER I - INTRODUCTION	
A. TWO MAJOR APPROACHES TO CLONING GENES	1
1. The 'Traditional' Approach	1
2. The Differential Approach	1
B. METHODS FOR CHARACTERIZING AND COMPARING mRNA POPULATIONS	5
1. Cot Curves and Reassociation Kinetics	5
2. Rot Curves and Hybridization Kinetics	6
3. Comparing mRNA Populations Using Hybridization Kinetics	8
4. Combining Subtraction with Kinetic Analysis	8
5. Saturation Analysis to Determine mRNA Complexity	9
6. Combining Subtraction with Saturation Analysis	10
7. Cascade Hybridization- Removing Common mRNA Sequences	11
8. Comparing mRNA Populations in Specialized Cell Types	13
9. Comparing mRNA Populations by cDNA Cloning	14
10. Purification of Specific cDNA Sequences Prior to the Advent of DNA Cloning	15
C. METHODS FOR CLONING DIFFERENTIAL GENES	17
1. Random Sampling and Screening	20
2. Differential Hybridization ('Plus-Minus')	21
3. Hybridization Competition	22
4. Subtractive Hybridization to Generate cDNA Probe	23
5. Subtractive Hybridization to Generate Subtracted Libraries	24
6. Subtracted cDNA Probe Used on Subtracted Libraries	25

7. Cloning Low Abundance Differential Genes- Which Method is Appropriate?..... 26

8. Other Methods for Cloning Differential Genes..... 27

D. DIFFERENTIATION OF LYMPHOCYTES DURING AN IMMUNE RESPONSE..... 29

1. Cells of the Immune System and Immune Recognition..... 29

2. Lymphocyte Activation..... 30

3. Lymphokines Secreted by Activated T-Helper Cells..... 33

4. Isolating New Lymphokines by cDNA Cloning..... 37

5. Tumor Models for Helper T-cell Activation..... 37

E. THESIS OBJECTIVE..... 38

CHAPTER II - MATERIALS AND METHODS

A. CELL CULTURE..... 40

1. Cell Culture Medium..... 40

2. Cell Lines..... 40

3. Reagents..... 41

4. Stimulation of Cell Lines to Produce Lymphokines..... 42

5. Biological Assay for IL2 Activity..... 42

B. ISOLATION AND CHARACTERIZATION OF RNA..... 43

1. Large Scale Isolation of Total Cellular and Poly A+ RNA..... 43

2. Small Scale Isolation of Cytoplasmic and Poly A+ RNA for Northern Analysis..... 44

3. Translation of Poly A+ RNA in Wheat Germ Extracts..... 45

4. Translation of Poly A+ RNA in *Xenopus laevis* Oocytes..... 46

5. RNA Gels and Northern Transfers..... 46

C. NUCLEIC ACID PROBES AND HYBRIDIZATION CONDITIONS.... 47

1. Oligonucleotide Probes..... 47

2. Labelling of cDNA Inserts and Other DNA Fragments..... 48

3. Prehybridization, Hybridization, and Washing Conditions for Blots... 48

D. PREPARATION OF cDNA LIBRARIES IN PHAGE LAMBDA GT10... 49

1. Preparation of Cloning Reagents..... 49

2. Testing Enzymes and Buffers for cDNA Cloning..... 54

3. Construction of cDNA Libraries..... 59

E. ISOLATION AND CHARACTERIZATION OF CLONES FROM cDNA LIBRARIES..... 64

1. Screening Libraries and Plaque Purification of Positive Clones..... 64

2. Preparing cDNA Inserts and Subcloning into Plasmid Vectors..... 65

F. ISOLATION OF PMA-INDUCED cDNA CLONES FROM EL4..... 67

1. Construction of Subtracted cDNA Library..... 67

2. Preparation of Subtracted cDNA Probe..... 70

3. Screening the Subtracted Library..... 72

G. CHARACTERIZATION OF PMA-INDUCED cDNA CLONES..... 73

1. Grouping Clones into Families, Subcloning, and Northern Analysis.. 73

2. Primary Sequence Analysis of PMA-Induced cDNA Clones..... 74

3. Expression of PMA-Induced cDNA Clones in COS Cells..... 80

H. CAP LABELLING AND HYBRIDIZATION TO MMTV SEQUENCES.. 81

CHAPTER III - INDUCTION OF INTERLEUKIN 2 MESSENGER RNA INHIBITED BY CYCLOSPORIN A

A. INTRODUCTION..... 83

B. CsA BLOCKS THE PRODUCTION OF IL2..... 84

C. CsA TREATMENT RESULTS IN A LOSS OF TRANSLATABLE IL2 mRNA..... 84

D. NO DETECTABLE IL2 mRNA IN CsA TREATED CELLS..... 88

E. CsA ALSO BLOCKS THE mRNA FOR GM-CSF IN EL4 CELLS..... 91

F. CONCLUSIONS..... 94

CHAPTER IV - ISOLATION OF cDNA CLONES ENCODING MOUSE IL2 AND MOUSE GM-CSF

A. INTRODUCTION..... 95

B. RESULTS..... 95

1. Frequency and Size of cDNA Inserts in the Library..... 95

2. Isolation of Murine IL2..... 97

3. Isolation of Murine GM-CSF..... 97

C. DISCUSSION..... 100

CHAPTER V - ISOLATION OF PMA-INDUCED cDNA CLONES FROM EL4 USING SUBTRACTIVE HYBRIDIZATION

A. INTRODUCTION..... 102

B. RESULTS AND DISCUSSION..... 105

1. Subtracted Library..... 105

2. Subtracted cDNA Probe..... 108

3. Probing the Subtracted Library..... 110

4. Grouping the cDNA Clones into Cross-hybridizing Families..... 113

5. Northern Analysis of Selected cDNA's..... 117

CHAPTER VI - CHARACTERIZATION OF PMA-INDUCED cDNA CLONES FROM EL4

A. INTRODUCTION..... 125

B. PATTERNS OF EXPRESSION OF PMA-INDUCED cDNA CLONES... 126

C. PRIMARY NUCLEOTIDE SEQUENCES OF THE PMA-INDUCED CLONES..... 136

D. DISCUSSION..... 145

CHAPTER VII - FURTHER CHARACTERIZATION OF THE 'CLONE 55' TRANSCRIPT

A. INTRODUCTION..... 149

B. RESULTS..... 150

1. Expression Pattern of the Clone 55 Transcript in Other Lymphoid Cell Lines..... 150

2. Expression of the Clone 55 Protein in COS Cells and Bioassay for Function..... 153

C. DISCUSSION..... 155

CHAPTER VIII - FURTHER CHARACTERIZATION OF THE SHORT PMA-INDUCIBLE MMTV TRANSCRIPT FROM EL4

A. INTRODUCTION..... 159

B. RESULTS..... 160

1. Sequence of the Most Abundant PMA-Induced Transcript in EL4..... 160

2. Expression Pattern of MMTV Related Transcripts in EL4..... 163

3. Determining the Initiation Site of the Short MMTV Transcript Using Cap Labelled RNA..... 165

4. Genomic Structure of the MMTV Provirus Present in EL4..... 165

C. DISCUSSION..... 168

CHAPTER IX - GENERAL DISCUSSION AND CONCLUSIONS..... 180

POSTSCRIPT..... 185

BIBLIOGRAPHY..... 186

LIST OF TABLES

		Page
1.	Examples Where a Differential Approach has been Used to Clone Genes....	18
2.	Preparation of Dideoxy Reaction Mixes for DNA Sequencing.....	78
3.	Effect of CsA on the Levels of Translatable IL2 mRNA in EL4.E1 Cells.....	87
4.	Translation of IL2 mRNA in Wheat Germ Extract.....	89
5.	Effect of CsA on the Levels of Translatable IL2 mRNA in Human Jurkat Cells.....	90
6.	Description of Synthetic Oligonucleotide Probes.....	96
7.	Summary of PMA-Induced cDNA Clones found in EL4.....	146
8.	Summary of Cell Lines Screened for Expression of the 'Clone 55' Transcript Using Cytodots.....	151
9.	Results of Bioassay for Lymphokine Activity in Transfected COS Cell Supernatants.....	154
10.	Bioassays in which COS Cell Supernatants Produced from Clone 55 have been Tested.....	156

LIST OF FIGURES

		Page
1.	The activation of T-helper lymphocytes results in the production of a number of lymphokine molecules	34
2.	Test of *Eco* RI linker ligation	58
3.	Effect of CsA on the production of IL2 by murine EL4.E1 cells	85
4.	Absence of IL2 mRNA in CsA-treated cells	92
5.	Induction of mRNA for IL2 and GM-CSF by PMA and supression by CsA in EL4.E1 cells	93
6.	Unselected EL4 cDNA library probed with 30 base oligonucleotide probe homologous to mouse IL2	98
7.	Rescreen of clone 17 using 34 base IL2 probe	99
8.	Final rescreen and plaque purification of mouse GM-CSF cDNA clone	99
9.	Procedure used to generate subtracted cDNA library	106
10.	Procedure used to generate subtracted cDNA probe	109
11.	Subtracted cDNA library probed with subtracted cDNA probe	111
12.	Rescreen and plaque purification of clone 25 cDNA	114
13.	Placing the PMA-induced cDNA clones into groups on the basis of cross-hybridization	115
14.	Clone 55 cDNA detects a PMA induced transcript in EL4.E1 cells	118
15.	Clone 154 cDNA detects a PMA induced transcript in EL4.E1 cells	119
16.	Clone 11 cDNA detects a very large number of PMA induced transcripts in EL4.E1 cells	121
17.	Clones 14, 15, and 22 detect a short PMA induced transcript in EL4.E1 cells	122
18.	Clone 32 cDNA detects a transcript which is present in both PMA treated and untreated EL4.E1 cells	123
19.	Expression pattern of the transcript detected by clone 12	127

20. Expression pattern of the transcript detected by clone 32..... 129

21. Expression pattern of the transcript detected by clone 25..... 130

22. Expression pattern of the transcript detected by clone 154..... 131

23. Expression pattern of the transcript detected by clone 14..... 132

24. Expression pattern of the transcript detected by clone 55..... 133

25. Expression pattern of the transcript detected by clone 55, using the cytodot method..... 135

26. Partial nucleotide sequence of clone 32 cDNA..... 137

27. Partial nucleotide sequence of clone 25 cDNA insert..... 139

28. Complete nucleotide sequence of clone 154 cDNA insert..... 140

29. Amino acid sequence and Hopp and Woods hydrophilicity plot of the protein coded by clone 154..... 141

30. Complete nucleotide sequence of clone 55 cDNA insert..... 143

31. Amino acid sequence and Hopp and Woods hydrophilicity plot of the protein coded by clone 55..... 144

32. Expression pattern of the transcript detected by clone 55 in the helper T cell line CH2.4..... 152

33. cDNA sequence of the predominant PMA-induced transcript in EL4..... 161

34. Northern blot analysis of poly A+ RNA from EL4, probed with clone 14.. 164

35. Cap labelled RNA demonstrates that the short MMTV transcripts initiate within the *env* region of the retrovirus..... 166

36. Southern blot analysis of EL4 genomic DNA cut with *Pst* I and probed with various segments of the MMTV provirus..... 169

37. Deletions found in the MMTV LTR for various murine T-cell leukemias... 175

38. Homologous nucleotide sequence found in the 5' flanking region of mouse IL2 and the *env* region of MMTV..... 178

ABBREVIATIONS AND DEFINITIONS

A-5m:	a gel media for exclusion chromatography (Bio-Rad).
Antibody:	an immunoglobulin molecule capable of binding in a specific fashion to a known molecule (antigen).
Antigen:	a molecule which can be recognized specifically by the immune system, the recognition of which stimulates an immune response.
AUG:	initiation codon for translation of mRNA.
B cell:	one of two major classes of lymphocytes, which have immunoglobulins on their surface, and respond to antigen by differentiating into antibody producing cells.
BCDF:	B cell differentiation factor.
BCGF:	B cell growth factor.
BRL:	Bethesda Research Laboratories (company).
BSA:	bovine serum albumin.
BSF-1:	B cell stimulating factor- one.
C3H:	an inbred strain of mice; also refers to a strain of MMTV isolated from these mice.
Cap:	the structure at the 5' end of eucaryotic mRNA added after transcription; the structure is $^{7Me}G^{5'}ppp^{5'}Np....$
Cascade hybridization:	repeated cycles of subtractive hybridization.
cDNA:	complementary deoxyribonucleic acid; DNA which has been copied from messenger RNA using the enzyme reverse transcriptase.
CHX:	cycloheximide, a compound which inhibits protein synthesis.
c I:	the gene which codes for the repressor protein of phage lambda; mutants in this gene cannot maintain lysogeny and always enter the lytic cycle to produce a clear plaque.
Con A:	concanavalin A- a plant lectin which is capable of stimulating T-cells to proliferate and differentiate.
COS cells:	a monkey fibroblast cell line which has been permanently transformed by an origin-defective mutant of SV40, and

	thus constantly expresses the SV40 large T antigen (Gluzman, 1981).
Cot:	concentration of DNA (moles of nucleotides per litre) multiplied by the time (in seconds) that the reassociation reaction has taken place. Standard reaction conditions are 25° C below the melting temperature, with neutral phosphate buffer at 0.12 M, monovalent cation at 0.18 M. At other salt concentrations and temperatures a correction factor is applied as described by Britten et al. (1974).
cpm:	counts per minute.
CsA:	cyclosporin A - a cyclic undecapeptide isolated from fungal extracts which has immunosuppressive properties.
CSF:	colony-stimulating factor; a growth factor which causes colonies to grow from bone marrow cells *in vitro* .
CTL:	cytotoxic T-lymphocyte.
Differential gene:	a DNA sequence which is expressed as mRNA at significantly different levels when two cells are compared.
DMSO:	dimethyl sulfoxide.
dNTP's:	the four common deoxynucleotide triphosphates (G,A,T,C)
ds cDNA:	double stranded cDNA.
DTT:	dithiothreitol.
EDTA:	ethylenediaminetetraacetic acid.
ED_{30}:	the amount of IL2 which causes 30% of the maximal response in a given bioassay is defined as 1 ED_{30} unit.
EL4.E1:	a subclone of the cell mouse cell line EL4 which was selected for its ability to produce high levels of IL2 after PMA stimulation; used for all the experiments described in this thesis (often called simply EL4).
env :	retroviral gene which codes for the envelope glycoproteins of the virion.
FCS:	fetal calf serum.
G50F:	sephadex G50-fine (Pharmacia); gel media for exclusion chromatography.
GM-CSF:	granulocyte macrophage colony stimulating factor.

GR:	an inbred strain of mice; also refers to a strain of mouse mammary tumor virus (MMTV) isolated from these mice.
HAP:	hydroxyapatite, used for chromatography to separate double stranded from single stranded nucleic acids.
hflA :	an *E. coli* gene, the product of which controls the stability of cII. cII is in turn a positive activator for cI synthesis, so that deletion of the *hflA* gene enhances cII stability, resulting in the stimulation of cI synthesis and a high frequency of lysogenization.
IL1, IL2, IL3, IL4:	interleukins 1, 2, 3, and 4.
JEH.3B:	subclone of the human T-leukemia cell line Jurkat, selected for its ability to produce IL2 in the presence of Con A and PMA.
kb:	kilobase or 1000 base pairs of double stranded DNA.
LAK:	lymphocyte activated killer cell.
lambda gt10:	a phage derived from lambda, which has a single *Eco* RI in the *c* I gene; often used as a cloning vector, particularly for cDNA libraries; can accept DNA fragments from 0 to ~7 kb.
LTR:	long terminal repeat; a segment of DNA which is repeated at each end of the proviral form of a retrovirus.
lymphokines:	soluble glycoproteins secreted by lymphocytes which have a biological effect on other cells (lymphocytes or otherwise).
mAP paper:	messenger affinity paper.
MCPM:	million counts per minute.
MHC:	major histocompatibility complex.
MMTV:	mouse mammary tumor virus; a retrovirus associated with mammary and other tumors in mice.
M_r:	molecular weight (relative).
mRNA:	messenger ribonucleic acid.
MTL2.8.2:	mouse T-cell line that proliferates in the presence of IL2, used to assay for IL2 (mouse or human); described by Bleackley *et al.* (1982).
NEN:	New England Nuclear (company).
NK:	natural killer cell.

NP-40:	nonionic detergent.
Ø-dil:	10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$.
PBS:	phosphate buffered saline, pH 7.2.
pcEXV3:	an eucaryotic cDNA expression vector; transcription of the inserted cDNA is from the SV40 early promoter, and an intron in the 5' untranslated region and a poly A addition signal are also included from SV40; also contains the SV40 origin of replication and therefore replicates episomally in COS cells (Miller *et al.*, 1985).
PDGF:	platelet derived growth factor.
Plus-minus, +/- :	a method of screening for sequences which are present in one set of mRNAs but absent from another, based on the differential hybridization of two separately prepared probes.
PMA:	phorbol-12-myristate-13-acetate; a phorbol diester compound capable of acting as a tumor promoting agent.
Pol I:	the enzyme DNA polymerase I (*E. coli*).
Poly A+ RNA:	cellular RNA enriched for mRNA by affinity chromatography over oligo-dT cellulose or mAP paper.
RHFM:	tissue culture medium containing RPMI 1640, hepes buffer, 2-mercaptoethanol, and fetal calf serum (Chapter II).
Rot:	concentration of RNA (moles of nucleotides per litre) multiplied by the time (in seconds) that the hybridization reaction (RNA with cDNA)has taken place. Standard reaction conditions are as described above for Cot.
RT:	the enzyme reverse transcriptase.
SDS:	sodium dodecyl sulphate.
SSC:	150 mM NaCl, 15 mM sodium citrate, pH 7.0.
ss cDNA:	single stranded cDNA.
Subtractive hybridization:	hybridization of mRNA to cDNA followed by separation of double stranded material from single stranded material.
SV40:	simian sarcoma virus; a much studied DNA tumor virus.
T3 complex:	three invariant polypeptide chains which are associated with the T-cell antigen receptor; recognized by a monoclonal antibody of the same name.

T4:	a glycoprotein expressed on the surface of helper T cells; recently renamed CD4.
T8:	a glycoprotein expressed on the surface of cytotoxic and supressor T cells; recently renamed CD8.
T-cell:	one of two major classes of lymphocytes, derived from the thymus and capable of recognizing antigens in combination with self-MHC molecules.
TdT:	the enzyme terminal transferase.
TE:	10 mM Tris-HCl, 1 mM EDTA.
Thalassemia:	a hereditary disease characterized by abnormally low syntheis or absence of one or more hemoglobin polypeptide chains; named according to the chain involved.
Tris-HCl:	tris (hydroxymethyl) aminomethane hydrochloride.
Vnx:	vanadyl sulfate-nucleoside complexes, a potent inhibitor of ribonucleases.

CHAPTER I

INTRODUCTION

A. TWO MAJOR APPROACHES TO CLONING GENES

1. *The 'Traditional' Approach*

Modern recombinant DNA technology has made it possible to isolate the genes which code for interesting biological proteins. This method of studying biological systems promises to be tremendously powerful and rewarding. The application of traditional biochemical ideas to the problem of isolating cloned genes has usually resulted in the following solution: a protein is discovered on the basis of a particular enzymatic or biological activity (or on the basis of an antibody which blocks a biological effect); the protein is purified and sequenced; and this sequence data is used to generate oligonucleotide probes (on the basis of the genetic code) which are in turn used to isolate the gene of interest. More recent developments in expression systems allow for the production of proteins directly from cloned DNA sequences (Noma *et al.*, 1986; Berg, 1981; Huynh *et al.*, 1985); this has obviated the need to purify and sequence the protein under investigation, since clones of interest can be identified and isolated directly on the basis of the proteins they produce. Both of these general approaches are predicated on the original discovery of the protein; that is, on the existence of some method of testing for the presence of the protein. In other words, the investigator must be aware of the existence of the protein before he can begin to clone the gene of interest.

2. *The Differential Approach*

In theory, a second general solution to the problem of isolating genes which code for biologically important proteins exists. The central idea underlying this approach is that of a differential gene. When two cells are compared, a differential gene is any gene

which is expressed (as RNA) at different levels in the two cells. A differential gene may be expressed at low levels in one cell, and much higher levels in the other (a relative differential), or it may be completely untranscribed in one cell, and highly expressed in the other (an absolute differential). In most cases, the cells to be compared contain exactly the same chromosomal (*i.e.* DNA) content, and the differential genes are those genes which are transcribed at different rates, and/or whose transcripts are processed or degraded at different rates in the two cells. In certain special cases the differential gene results because (at the DNA level) the gene is absent in one cell type and present in the other (a genomic differential).

Differential genes include all traditionally described inducible genes, in both procaryotic and eucaryotic cells. The general case is one in which a cell with phenotype **a** (which contains a set of RNAs {**A**}) is placed under inducing conditions (*e.g.* by addition of a metabolite or a hormone) which causes the cell to change the expressed set of RNAs to {**A'**} and so alter its phenotype to **a'**. The induced genes then fall in the set {**A'-A**}. In complex multicellular organisms, differential genes also include those genes which are expressed at different levels in different cell types. In the general case, a cell with phenotype **a** (*e.g.* a liver cell) expresses a set of mRNAs {**A**}, and this can be compared to a cell with a different phenotype **b** (*e.g.* a fibroblast) which expresses the set of mRNAs {**B**}. The set of differential genes then includes {**B-A**} (fibroblast compared to hepatocyte) and {**A-B**} (hepatocyte compared to fibroblast). In this description the difference between cell **a'** and cell **b** is simply that we know how to derive **a'** from **a**, and can do so at will, whereas we do not understand how cell **b** is derived from cell **a**. Although cells **a** and **b** originate from the same cell and usually have an identical genotype, they follow different differentiative pathways during embryological development. In species which reproduce sexually, differential genes also include those genes which reside only on the Y chromosome and are expressed

only in males.

In general, differential genes code for biologically important (and therefore interesting) proteins. This is because differential genes (and particularly absolute differential genes) are the very genes which code for the proteins required by the cell to carry out its specialized function. (Some authors have labelled such proteins 'luxury' proteins, to contrast them with 'housekeeping' proteins which are made by and required for the maintenence of all cells[Lewin, 1980].) The nature of DNA (and RNA) makes it possible to clone differential genes directly, and examples of differential genes and the strategies used to clone them will be described below.

The differential approach to the problem of cloning genes can be compared with the 'traditional' approach described in the first paragraph above. The two appear to be almost the inverse of one another. The latter approach requires that the investigator be able to assay for the protein of interest, and requires *a priori* awareness of the existence of the protein and its possible function. The differential gene approach does not necessarily require prior knowledge of the properties of (or even the existence of) a particular protein, but requires that the investigator chooses for comparison two cells where the differential genes are significant. When this results in the discovery of new proteins, their functions remain unknown unless they are homologous to proteins of known functions. Clearly the two approaches are complementary.

In a general way the differential approach to cloning genes has similarities to the mathematical solution of a problem in classical physics. In physics one considers a physical problem, and by making certain assumptions translates the problem into mathematical language (usually in the form of a differential equation). The equation is solved by using purely mathematical operations, and the mathematical solution must then be interpreted to determine what it means in terms of the physical situation. Often this last step is the most difficult one. In comparison, with differential cloning one

makes certain assumptions about what constitutes an important biological difference between two cells. Without further consideration of the biology of the system, differential genes can be isolated and characterized. The solutions to the biological differential (*i.e.* the sequences of the differential genes) must then be interpreted to determine what they mean in terms of the biological system under consideration. This last step will almost always be the most difficult for molecular biologists, and will require new approaches to the problem of determining the function of an unknown protein (such as 'reverse genetics' where a gene is added to a system, or 'anti-expression' where the transcripts of a gene are effectively subtracted from a system by using anti-sense RNA).

In most interesting cases differential genes are manifest at the level of RNA, and the two cells under consideration are identical or nearly identical with respect to their genomic DNA. To isolate differential genes, therefore, the primary starting material is necessarily RNA or cDNA (DNA copied directly from the RNA and therefore containing the same information as the original RNA). Strictly speaking, a differential gene is one which is contained in the nuclear DNA of the cell, and is therefore obtained from a genomic library constructed from chromosomal DNA. However, for a number of reasons differential genes are often first isolated from cDNA libraries. Whereas genomic clones often contain DNA sequences from several adjacent genes (only some of which may be differential genes), each clone in a cDNA library contains sequences from only one gene, and this simplifies the analysis. With a cDNA library fewer clones need to be screened, and this saves time and materials. Furthermore, the sequence of the cDNA clone is required to determine the amino acid sequence of the protein coded by the differential gene. For the purposes of this work, therefore, differential genes will be considered cloned if they are isolated at either the cDNA level or the genomic level, since the isolation of either one of these types of clones can lead to the isolation of

the other.

Since differential genes are primarily manifest at the level of RNA, it is worthwhile to consider the methods which have been used to characterize the RNA species in a given cell, and to measure the differences in RNA between different cell types. This topic also serves as a useful introduction to the ideas which follow.

B. METHODS FOR CHARACTERIZING AND COMPARING mRNA POPULATIONS

1. *Cot Curves and Reassociation Kinetics*

The reassociation of two complementary strands of nucleic acid (DNA or RNA) depends upon random collision and follows second order kinetics, providing that the complementary strands are at equivalent concentrations (Lewin, 1980). The solution of the corresponding rate equation demonstrates that the reassociation of any given double stranded nucleic acid can be characterized by a number $C_0t_{1/2}$, which is the product of the initial concentration of the nucleic acid in the hybridization mixture (in nucleotide moles/litre) multiplied by the time (in seconds) required for the reaction to be half complete. The reassociation of nucleic acids is usually followed by drawing a 'Cot' curve, which plots the fraction of the nucleic acid which has formed double stranded material (*i.e.* the progress of the hybridization reaction) against the log of the Cot (product of C_0 and time). When the Cot curves for DNA reassociation were compared for various organisms, it became clear that organisms with smaller genome sizes had smaller $C_0t_{1/2}$ values. This is a consequence of the fact that it is the concentration of each reassociating sequence that determines its rate of renaturation. At a given DNA concentration, if the genome is short it will be represented many times and therefore be at a higher concentration (and reassociate at a faster rate), whereas if a genome is much

longer it will be represented correspondingly fewer times and therefore be at a lower concentration (and reassociate more slowly). The idea of the 'length of the genome' and its inverse relationship to $C_0t_{1/2}$ values can be generalized to any case where the rehybridization kinetics of nucleic acids are studied. In an uncharacterized sample of nucleic acids, the $C_0t_{1/2}$ of the reaction reflects the total length (also called the complexity) of all the different sequences present, and this value is usually expressed in base pairs (Lewin, 1980).

2. *Rot Curves and Hybridization Kinetics*

The complexity and composition of a messenger RNA population can also be investigated by studying the kinetics of rehybridization. In this case the mRNA is primed with oligo-dT and copied into cDNA using the enzyme reverse transcriptase (usually a radioactive nucleotide is incorporated into the cDNA). The double stranded DNA-RNA hybrid molecules are then melted and allowed to rehybridize in an excess of the same RNA. Since one of the hybridizing components is now in great excess, its concentration is essentially constant and the rehybridization reaction now follows pseudo first-order kinetics (Hood *et al.*, 1975). The progress of the reaction is followed by monitoring the appearance of double stranded radioactive material, usually by hydroxyapatite chromatography (hydroxyapatite retains only double stranded nucleic acids, and allows single stranded material to pass through) or by S1 nuclease treatment (S1 nuclease degrades only single stranded nucleic acids, and leaves double stranded material intact). The data from such an experiment is displayed in the form of a Rot curve, where the log of Rot (initial concentration of RNA multiplied by the time in seconds) is plotted versus the percentage of hybridization. For typical eucaryotic cells the shape of the resulting Rot curve is complex, but it can usually be resolved (by computer analysis) into three components, each representing a certain fraction of the

total counts hybridized. For each component a characteristic $R_o t_{1/2}$ can be determined based on the value of Rot at which half of the material for that component was hybridized. The value of $R_o t_{1/2}$ for each component can then be used to make an estimate of the complexity of that sub-population of RNA, based on comparison with the $R_o t_{1/2}$ of a standard of known complexity.

For example, if 50% of the mRNAs in a cell consist of a single sequence 1000 bases in length, this sequence will be present at a high concentration and will hybridize rapidly and move 50% of the cDNA counts into double stranded material. The $R_o t_{1/2}$ of this component will be small and will correspond to a complexity of 1000 base pairs. The remaining RNA sequences will be present at lower concentrations, and will therefore hybridize more slowly and yield correspondingly larger $R_o t_{1/2}$ values. These values in turn can be used to estimate the total complexity of the RNA sub-populations representing the intermediate and slow reacting components of the Rot curve. If an assumption is made about the average size of the mRNAs, the complexity can be used to estimate the total number of different mRNAs present in each component. Since the proportion of the total mRNA present in each component is known, the average number of copies of the messengers in each component can be estimated if the total amount of mRNA in the cell is known (Lewin, 1980). This method of analyzing the kinetics of hybridization of mRNA to its cDNA was first used by Bishop *et al.* (1974) to characterize HeLa cell mRNA. These and subsequent studies have shown that the 'typical' mammalian cell contains on the order of 4 different highly abundant mRNAs (≈12,000 copies/cell), 500 different moderately abundant mRNAs (≈300 copies/cell), and 11,000 different low abundance mRNAs (≈15 copies/cell) (Alberts *et al.*, 1983). However, these figures vary considerably between different cell types.

3. *Comparing mRNA Populations Using Hybridization Kinetics*

Kinetic studies of the hybridization of mRNA to its cDNA are useful in analyzing the messages present in a given cell, but they do not allow for the comparison of the mRNAs present in two different cells. However, by hybridizing the cDNA from one cell to an excess of mRNA from another cell, general comparisons of the abundance classes of the mRNAs which are common to the two cells can be made. Furthermore, the total amount of hybridization at high Rot values gives an estimate of the total amount of overlap that is present between the mRNA sets present in two different cells. Hastie and Bishop (1976) made such a comparison among the sets of mRNAs expressed in mouse liver, kidney, and brain, and concluded that a high proportion of the total sequences were common among the three tissues. Furthermore, their analysis showed that the abundant mRNA of the liver (for example) is different from the abundant mRNA of the kidney and brain.

4. *Combining Subtraction with Kinetic Analysis*

The amount of information which can be obtained by analysis of hybridization kinetics using cDNA of one cell and an excess of mRNA from another is limited. It is impossible to tell, for example, if the sequences which are shared are in the same or different abundance classes. More information can be gained if a subtractive analysis is performed in combination with further kinetic analysis. This technique depends on isolating cDNA of a certain abundance class from a given cell, and then hybridizing it with an excess of mRNA from a second cell and determining in which abundance class the common sequences are found in the second cell type. This type of analysis was also undertaken by Hastie and Bishop (1976). The abundant cDNA of mouse liver, for example, was isolated by hybridizing the cDNA to its homologous mRNA to a low Rot, and then separating the double stranded material from single stranded by hydroxyapatite chromatography (in this short time only the highly abundant sequences

would have hybridized to form double stranded material). The cDNA can be isolated from the hybrid by base hydrolysis, and then used to hybridize with an excess of mRNA from kidney or brain. These analyses showed that the mRNAs of the high abundance class in liver were actually found in the low abundance mRNA species in kidney and brain. A similar analysis was done for middle and low abundance mRNAs from kidney, and it was shown that kidney and liver shared some middle abundance sequences, and that all three tissues share a great many low abundance mRNAs. It was suggested by the authors (Hastie and Bishop, 1976) that the majority of the low abundance mRNAs may be involved in 'housekeeping' functions, and that it is primarily the most abundant class of mRNAs which is responsible for the specialized functions of a cell. Their analysis does not prove, however, that certain middle and low abundance mRNAs may not also play a critical role in the specialized fuctions of a cell. The mRNAs coding for the T-cell antigen receptor (Hedrick *et al.*, 1984a) provides an example of the latter case (see below).

Kinetic analysis is complicated by the fact that similar but non-identical sequences may still be able to hybridize with each other, but at a rate which is lower than would be predicted by their concentration if they were identical. Kinetic analysis and subtractive-kinetic analysis such as those described by Hastie and Bishop are inherently limited by the fact that only general, overall properties of mRNA sets can be assessed, but an understanding of the techniques used and the results obtained (and their interpretation) are useful for the purposes of this thesis.

5. *Saturation Analysis to Determine mRNA Complexity*

The total compexity of mRNA present in a cell can also be measured by a saturation analysis (as opposed to the kinetic analysis described above). DNA is extracted from a cell and the total complexity of the nonrepetitive fraction of the DNA is determined by a Cot analysis (described in the first paragraph of this section). The

nonrepetitive fraction of the DNA is then isolated by melting and reannealing to a low Cot value (at which only the highly abundant, repetitive fraction of the DNA will have formed double stranded DNA) and then separating the double stranded material from the single stranded material by hydroxyapatite chromatography. This process can be repeated several times to obtain a preparation of DNA that contains only the nonrepetitive fraction of the DNA, the complexity of which is known. This material can be labelled and then hybridized to an excess of cellular RNA. The highly abundant mRNAs will rapidly hybridize to their complementary DNAs, and the reaction will then be driven by the mass of low abundance mRNAs present in the population. (Again this is a pseudo-first order reaction, but the effective RNA concentration driving the reaction is lower than the measured concentration since the highly abundant RNAs do not drive the reaction.) If the hybridization is allowed to continue to very high Rot values, a plateau is reached beyond which no further hybridization will occur. The amount of the DNA which is hybridized at this saturation point can be determined (it is typically a few percent of the total DNA), and since the complexity of the total nonrepetitive DNA is known, the complexity of the RNA population that hybridizes to it can be calculated. Saturation analysis can also be used to compare total differences in mRNAs from different cells by performing a double saturation analysis. Labelled DNA is hybridized with an excess of a mixture of RNA derived from two different cells, and the saturation level is compared with that obtained with either sample of mRNA alone. An increase in the double saturation above the level of either tissue alone identifies sequences which are not shared between the two RNA populations, whereas a reduction below the sum of the two separate saturation levels is due to sequences which are common to both RNA populations (Lewin, 1980).

6. *Combining Subtraction with Saturation Analysis*

A variation on saturation analyis, coupled with a subtractive technique, has been

used to determine overlaps between mRNA sets present in different cells (this has been called exhaustion analysis [Lewin, 1980]). This type of analysis was done by Galau *et al.* (1976) to compare the sets of mRNAs present in sea urchins at various stages of embryonic development, as well as in adult tissues. For example, excess gastrula mRNA was hybridized to labelled, nonrepetitive DNA to a high Rot, and double and single stranded species separated on hydroxyapatite chromatography. This process was repeated, and the end result was two purified populations of DNA: one containing sequences able to hybridize to the mRNAs present in cells of the gastrula stage ('gastrula mDNA'); and one DNA population containing no sequences in common with gastrula mRNA ('null mDNA'). Samples of these two labelled populations of DNA were then hybridized to excess mRNA from cells of various stages of embryonic development, and the complexity of the mRNA capable of hybridizing to the gastrula mDNA and the null DNA determined. For example, the mRNAs present in an oocyte contained a subset of equal complexity to the gastrula mDNA, and another very large subset of approximately equal complexity capable of hybridizing to the null mDNA. The oocyte therefore contained virtually the same set of mRNAs present in the gastrula, plus a large additional set of mRNAs which were not found in the gastrula. A similar type of analysis was done to compare the mRNAs expressed in various tissues of the mature tobacco plant (Kamalay and Goldberg, 1980). In general this type of analysis is useful in giving an overall picture of the size of the sets of mRNAs which are common to different cell types, as well as the size of the mRNA sets which are not shared by different cells (*i.e.* the size of the differential gene set). However, the abundance class of the various common and differential mRNAs cannot be determined without further kinetic analysis.

7. *Cascade Hybridization- Removing Common mRNA Sequences*

Another important method for comparing the sets of mRNAs present in different

cells is described by Timberlake (1980). During differentiation of the fungus *Aspergillus nidulans* the cells change from a vegetative state of growth to one where specialized spore-forming organs (conidiophores and conidia) and spores are produced. A kinetic analyis of cDNA-mRNA hybridizations showed that the vegetative cells express mRNAs of three abundance classes, and furthermore that sporulating cultures contain a set of mRNAs which is not expressed by cells in the vegetative state.

To further assess the set of mRNAs expressed by the differentiating spore-forming organs but not by the vegetative cells, cDNA representative of this differential set of mRNAs was isolated using the method of cascade hybridization (Timberlake, 1980). Labelled cDNA was transcribed from mRNA of sporulating cultures, and the mRNA removed by base hydrolysis. The labelled cDNA was hybridized to a 20-fold excess of mRNA from vegetative cells, and single stranded material was separated from double stranded material by hydroxyapatite chromatography. This process was repeated twice more, using a 50- and then 100-fold excess of vegetative mRNA. The single stranded cDNA remaining was completely depleted of any sequences found in the vegetative cells, and thus contained only sequences which were found in sporulating cultures but not in vegetative cultures. From this sample the 'spore-specific' sequences were isolated by hybridizing to mRNA from spores, and then collecting the double stranded material. From the single stranded material still remaining the 'non-spore, development specific' sequences were isolated by hybridizing to mRNA from sporulating cultures and then collecting double stranded material. These final two positive selection steps were necessary since a portion of the single stranded material which remains after subtractive hybridization of the vegetative sequences is simply unreactive with any mRNA and would confuse the subsequent analysis. Furthermore sporulating cultures contain spores as well as conidia and conidiophores, and sequences from all three cell types are represented in the subtracted

material. Since pure spores can be isolated at a slightly later stage, mRNA from these can be used to separate the spore-specific sequences from the remainder of the development specific sequences. The spore-specific and non-spore, development specific sequences differ from the null mDNA sequences described above in that the former contain just those sequences which are present in one developmental stage and absent from the previous one, whereas the null mDNA contains all of the possible sequences which are absent from one developmental stage. Analysis of the hybridization kinetics of excess homologous mRNA with the developmental specific sequences isolated demonstrated that there were about 300 spore-specific, and 1000 nonspore development specific mRNAs. Furthermore, in each case the cell-specific mRNAs fell into three abundance classes, with the greatest number of sequences present in the low abundance class (Timberlake, 1980). These results do not disagree with those of Hastie and Bishop (1976), but they do point out that mRNAs which are unique to one cell type and probably involved in the specialized function of the cell can fall into the low abundance class of mRNAs, as well as into the highest abundance class (see above).

8. *Comparing mRNA Populations in Specialized Cell Types*

Most of the methods used to characterize the mRNA set expressed by a given cell and to compare mRNA sets expressed by different cells were originated by investigators interested in general problems of developmental biology. However, the methods are general and can be applied to comparisons between mature cells which are closely related developmentally and functionally. In general, the more closely related the two cells are, the smaller is the difference between the sets of mRNAs expressed. Crampton *et al.* (1980) used homologous and heterologous mRNA-cDNA hybridizations to show that human fibroblasts and lymphocytes differed in about 25% of the mRNA sequences expressed. Davis *et al.* (1982) used the method of cascade/

subtractive hybridization to assess differences in mRNAs expressed by B and T lymphocytes, and showed that the difference amounted to about 2% of the mRNAs expressed. They also showed that mRNA sets expressed by different T-cell lines differed by 0.2-1.0% of the total sequences. In another set of experiments Hedrick *et al.* (1984a) used cascade/subtractive hybridization to show that the set of mRNAs coding for membrane bound or secreted proteins in T-helper lymophocytes (*i.e.* membrane bound polysomal fraction of the mRNA) differed from the mRNAs expressed in B lympohocytes by ≤0.5% of the total sequences. Rot analysis of these subtracted cDNA sequences demonstrated that the mRNAs specific to T-cells were derived from both high and low abundance classes, and that the largest group (consisting of about 70 different sequences) was from the low abundance class. Presumably mRNAs from this low abundance class represent sequences which are important to the specialized function of T-helper lymphocytes. This was later shown to be the case for at least one of the mRNAs in this class, which codes for the ß-chain of the T-cell antigen receptor (Hedrick *et al.*, 1984b).

All of the methods described above for comparing mRNA sets expressed by different cells are statistical in nature. They yield general information on the size and complexity of the sets of mRNAs which are differentially expressed between two cells, as well as the sets of mRNAs which are common to different cells. In general, they do not give information about specific sequences which are differentially expressed.

9. *Comparing mRNA Populations by cDNA Cloning*

If mRNA sets which are very similar in composition are compared, it is possible to imagine that the small number of differential mRNAs could be individually identified and then actually counted to determine the differences between the two RNA sets. With modern methods of DNA cloning (and particularly cDNA cloning), it is possible to isolate and identify sequences corresponding to individual mRNA transcripts. Cloning

differential genes thus represents another approach to the problem of determining the differences in sets of mRNAs expressed by different cells. This approach has the advantage that the nature of the specific sequences which are differentially expressed can be determined. However, to characterize exhaustively the differences in the sets of mRNAs expressed by two cells requires that the total set of differential genes be cloned, and this has not generally been possible with existing technologies. For the most part, the differential genes which have been cloned by developmental biologists represent a 'hit and miss' sample of the high and intermediate abundance mRNAs which are differentially expressed between developmental stages (see below). These experiments fall far short of cloning the total set of differential genes, and in particular lack differential genes from the low abundance class of mRNAs, some of which probably code for products important in the control of the differentiation process itself. In most of the other cases where differential genes have been cloned, the investigators have used a differential approach to clone a particular gene (often representing a low abundance mRNA; for example, interferon). These experiments do not attempt to clone the total set of differential genes, but the methods they use may have some relevance for the solution of the more general problem. Certainly any method for isolating clones representing the total set of differential genes must be able to detect sequences representing mRNAs from the lowest abundance class. The various methods which have been used and the results obtained are discussed in I. C below.

10. *Purification of Specific cDNA Sequences Prior to the Advent of DNA Cloning*

Prior to the invention of DNA cloning, sequences representing individual differential genes were isolated in a few special cases, and these are interesting if only to point out that the idea of and the methods for isolating differential sequences are not new. In all cases a single biological differential had been defined between the two systems under investigation. Alt *et al.* (1978) isolated cDNA sequences complementary

to dihydrofolate reductase mRNA by exhaustively hybridizing cDNA from a methotrexate resistant cell line to an excess of mRNA from the original methotrexate sensitive cell line. The non-hybridized material (representing the sequences differentially expressed in the resistant cell line [a relative differential]) was recovered by hydroxyapatite chromatography, and positively selected by hybridizing to mRNA from the resistant cell line. The final purified cDNA was shown by rehybridization kinetics to consist of a single component, which hybridized to dihydrofolate reductase mRNA. The purified sequences were used to demonstrate that the dihydrofolate reductase genes were amplified selectively in methotrexate-resistant cell lines. Stehelin *et al.* (1976) were able to isolate cDNA complementary to the transforming sequences of the avian sarcoma virus by hybridizing cDNA from a wild-type virus to RNA from a transformation defective strain of the virus which had deleted the oncogenic sequences. In this case two cycles of subtractive hybridization followed by hydroxyapatite chromatography were used, followed by a final positive selection step using RNA from the wild-type virus. Ramirez *et al.* (1975) used a similar strategy to purify α and β globin cDNAs by hybridizing total human globin cDNA to mRNA from the liver of a hydrops fetalis (α thalassemia); the β cDNA hybridized to the hydrops fetalis mRNA while the α cDNA remained single stranded, and these were separated on hydroxyapatite. Alt *et al.* (1979) purified cDNA sequences specific for immunoglobin α, λ and μ heavy chains by the method of 'subtractive hybridization.' The generalized purification scheme involved hybridizing cDNA from myeloma 1 (expressing heavy chain A and light chain X) to an excess of mRNA from myeloma 2 (expressing heavy chain B and light chain X) along with mRNA from a fibroblast cell line (which presumably expresses the same set of 'housekeeping' mRNAs as do the myelomas).

The non-hybridized cDNA, now highly enriched for heavy chain A specific sequences, was collected by hydroxyapatite chromatography. As a final step, the A specific sequences were positively selected for by hybridization to mRNA from myeloma 1 and isolation of double stranded material. The purified heavy-chain specific cDNA sequences were used to study the expression of the various heavy chains in a number of myeloma cell lines, by analysis of cDNA-mRNA hybridization kinetics.

C. METHODS FOR CLONING DIFFERENTIAL GENES

The 'typical mammalian cell' contains mRNAs which fall into three abundance classes, as described above. The relative frequency of clones representative of the various abundance classes of mRNA can be calculated, at least for an 'ideal' cDNA library (which contains all of the sequences present in the original mRNA, and in the same relative proportion). Assuming the cell contains a total of 360,000 mRNA molecules, clones containing sequences from the high abundance class (12,000 copies per cell) would be present at a frequency of 3.3% (1:30). Clones containing sequences from the middle abundance class (300 copies per cell) would be present at a frequency of 0.08% (1:1200). Clones containing sequences from the low abundance class (15 copies per cell) would be present at a frequency of 0.004% (1:24,000). Extremely rare transcripts may be present at 5 mRNA copies per cell, and would be found in the library at a frequency of 0.0014% (1:72,000). These figures are useful when considering the frequencies of clones which can be detected by the various methods used to clone differential genes.

Table 1 gives a summary of a large number of published reports where a differential approach has been used to clone various genes. The examples chosen are listed in relatively chronological order as they appeared in the literature. The list attempts to be quite comprehensive for the first few years; thereafter, examples have

TABLE 1

EXAMPLES WHERE A DIFFERENTIAL APPROACH HAS BEEN USED TO CLONE GENES

Biological differential	Method	Gene(s) cloned	Abundance	Reference
Yeast metabolizing galactose vs. glucose	+/-, cDNA probe	5 galactose induced genes (genomic clones)	0.01%	St. John and Davis, 1979
Poly (I):(C) induced fibroblasts vs. uninduced fibroblasts	Subtractive* hybridization	Fibroblast interferon (fl) cDNA clone	0.03%	Taniguchi *et al.*, 1979
Developing slime mold (6 hr.) vs. vegetative cells	+/-, mRNA probe	23 developmentally regulated cDNAs	0.1%	Williams and Lloyd, 1979
Aspergillus conidiating cells vs. vegetative cells (mycelia)	Subtractive hybridization	350 clones induced during conidiation	≥0.01%	Zimmermann *et al.*, 1980
Developing slime mold (6 hr., 15 hr.) vs. vegetative cells	+/-, mRNA probe	8 develpmentally regulated cDNAs	0.1%	Rowekamp & Firtel, 1980
Gastrula vs. Tadpole in developing Xenopus embryos	+/-, cDNA probe	many stage-specific cDNA sequences	0.06%	Dworkin and Dawid, 1980
Several stages in the embryonic development of sea urchins	+/-, cDNA probe	several stage-specific cDNA sequences	0.1%	Laskey *et al.*, 1980
Human lymphocytes vs. human fibroblasts	+/-, mRNA probe	5 cDNAs differentially expressed in lymphocytes	0.2%	Crampton *et al.*, 1980
B-hybridoma (k,μ,J) vs. L-cell also +ive selection with J	Subtractive hybridization	cDNA for J-chain of IgM pentamer	0.2-0.3%	Mather *et al.*, 1981
Several developmental stages of Dictyostelium	Hybridization competition	several stage-specific cDNA sequences	N.D.	Mangiarotti *et al.*, 1981
Enterotoxin induced lympho-cytes vs. resting cells	+/-, cDNA probe*	human gamma interferon cDNA	0.04%	Gray *et al.*, 1982
Embryonic Myoblasts vs. differentiated myofibers	+/-, cDNA probe	actin, tropomyosin, myosins, troponins	~0.1%	Hastings and Emerson, 1982
SV40 transformed 3T3 cells vs. non-transformed cells	+/-, mRNA probe	cDNA sequences specific to tranformed cells	0.25%	Schultzbank *et al.*, 1982
EGF treated mouse embryo cells vs. untreated cells	+/-, mRNA probe	3 gemonic clones all related to retrovirus VL30	N.D.	Foster *et al.*, 1982
Xenopus embryonic gastrula vs. oocyte	Subtractive cloning	many gastrula-specific cDNA sequences	0.001%	Sargent and Dawin, 1983
Schizophyllum metabolizing cellulose vs. glucose	Subtractive cloning	several genes involved in cellulose breakdown	N.D.	Seligy *et al.*, 1983
PDGF treated 3T3 cells vs. untreated cells	+/-, cDNA probe	5 cDNA sequences induced by PDGF	0.1%	Cochran *et al.*, 1983
Proliferating 3T3cells(+serum) vs confluent cells (low serum)	+/-, cDNA probe	several 'serum' induced cDNA sequences	0.1%	Linzer and Nathans, 1983

N.D.: not determined; *mRNA was size selected to enrich sequences of interest prior to subtraction and cloning
Unless otherwise stated the method of subtractive hybridization refers to the preparation of cDNA probes.

TABLE 1 (Cont.)

EXAMPLES WHERE A DIFFERENTIAL APPROACH HAS BEEN USED TO CLONE GENES

Biological differential	Method	Gene(s) cloned	Abundance	Reference
Rat brain vs. rat liver and kidney	northern blots of each clone	rat brain specific cDNA sequences	0.01%	Milner and Sutcliffe,1983
G1 baby hampster kidney cells vs. G0 cells	+/-, cDNA probe	5 cDNA sequences, cell cycle specific (G1)	N.D.	Hirschhorn *et al.*, 1984
Mouse B-lymphocyte cell line vs. T-cell line	Subtractive hybrid. (probe & library)	genomic& cDNA clones for mouse I-A_a	0.01%	Davis *et al.*, 1984
Mouse T-lymphocyte# cell line vs. B cell line	Subtractive hybrid. (probe & library)	ß-chain cDNA for T-cell antigen receptor (mouse)	0.001%	Hedrick *et al.* 1984a
Human T-lymphoma vs. B cell line	+/-, cDNA probe	ß-chain cDNA for T-cell antigen receptor (human)	N.D.	Yanagi *et al.*, 1984
Mouse T-lymphocyte# cell line vs. B cell line	Subtractive hybrid. (probe & library)	γ and α chain cDNAs for T-cell antigen receptor	N.D.	Saito *et al.*, 1984a,1984b
T-helper hybridoma 1 vs T-helper lines 2 and 3'	Subtractive hybridization	α chain cDNA for T-cell antigen receptor (mouse)	0.001%	Chien et al., 1984
Transfected L-cells expressing T8 vs. untransfected L-cells	Subtractive hybridization	lymphocyte surface marker T8	N.D.	Kavathas,1984 Littman, 1985
Male mouse chromosomal DNA vs. female DNA	Deletion enrichment & cloning	genomic clones specific to the Y-chromosome	N.D.	Lamal and Palmer, 1984
Yeast induced with mating phermone vs. uninduced cells	nascent RNAs labelled with thiouridine	5-genomic clones with phermone-induced sequences	N.D.	Stetler and Thorner,1984
Transfected L-cells expressing T4 vs. untransfected L-cells	Subtractive hybridization	lymphocyte surface marker T4	N.D.	Maddon *et al.*, 1985
Normal human cells vs. cells with a chromosomal deletion	Deletion enrichment & cloning	clones for sequences deleted in the mutant	N.D.	Kunkel *et al.*, 1985
Mutant aspergillus vs. wild type cells	+/-, cDNA probe	genomic sequences overproduced in the mutant	N.D.	Atkinson *et al.* 1985
Conidiating neurospora vs. vegetative cells	Subtractive hybridization	genomic sequences induced during conidiation	0.1%	Berlin & Yanofsky, 1985
Cytotoxic T-lymphocyte line vs. helper T-cell & thymocyte	+/-, cDNA probe	2 CTL specific cDNA sequences	0.5%	Lobe *et al.*, 1986
Cytotoxic T-lymphocyte cell line vs. thymic lymphoma	hybridization competition	CTL specific cDNA sequence	N.D.	Gershenfeld& Weissmann, 1986

N.D.: not determined; #membrane-bound polysomal fraction of T-cell RNA was used to enrich for secreted and membrane-bound sequences; 'this strategy depends on differences between the sequences of the variable regions of the various alpha chains of the T-cell receptor.
Unless otherwise stated the method of subtractive hybridization refers to the preparation of cDNA probes.

been included either because they represent interesting classes of genes (mitogen induced, growth related, or immunologically important genes), or because they illustrate new methodological approaches to the problem of cloning differential genes. A comprehensive list of all the genes which have been cloned in the past three years using a differential approach would probably include several hundred entries, so the list in Table 1 cannot be considered complete!

1. *Random Sampling and Screening*

The simplest approach to cloning differential genes involves random sampling and screening. Clones (usually cDNA clones) are simply picked at random, and their DNA inserts isolated and labelled. The labelled inserts are used to probe northern blots containing mRNA from the cells to be compared. This is the approach taken by Milner and Sutcliffe (1983) to clone cDNA sequences differentially expressed in rat brains. It is worthwhile considering for a moment what one is likely to find if clones are simply picked at random from an 'ideal' cDNA library made from a 'typical' mammalian cell. About 13% of all the mRNA molecules fall into the high abundance class ([12,000 molecules/cell] x [4 different molecules]= 48,000 molecules or 13% of the total of 360,000 molecules/cell; see above), about 41% into the moderate abundance class, and about 46% into the low abundance class. Most of the clones picked will therefore represent sequences in the moderate or low abundance classes, with about half falling into each class. The chance of picking a differential gene is directly proportional to the size of the mRNA differential between the two cells under consideration. If two cells differ by a large proportion of their expressed sequences (say 25%), then there is a large chance (1:4) that randomly chosen clones will contain differential sequences. Milner and Sutcliffe were able to pick at random a large number of brain-specific genes because they compared cells which have a very large differential (>30%). A large number of their sequences fall into the moderate to low abundance classes simply for

statistical reasons (brain has a higher proportion of mRNAs in the low abundance class than does a 'typical' cell, so these are even more likely to be picked than the figures above would indicate). One of the technical difficulties that can be encountered in a random analysis of a cDNA library is that cDNA sequences of very low abundance may not be detectable on northern analysis. Milner and Sutcliffe found that 26% of their clones failed to detect mRNA in any of the three tissues (brain, liver, kidney) examined, although they gave bands on southern analysis, and presumably represented transcripts which are present in brain (but are estimated by the authors to be present at <0.01% of the transcripts). Whether or not they are present only in brain cannot be determined. Others have suggested that even the lowest abundance mRNAs can be detected on northern analysis (down to 0.001% of the transcripts), so that Milner and Sutcliffe's failure to detect mRNA corresponding to low abundance cDNAs may have been due to a technical problem.

2. *Differential Hybridization ('Plus-Minus')*

Random screening is very likely to find differential genes if cells with a large biological differential are compared. However, if very similar cells are compared (say with 2% difference in the sequences expressed) the problem of finding differential genes becomes much more difficult and demands a more directed strategy. The simplest such strategy is differential hybridization (also known as 'plus-minus' screening; called +/- in Table 1). A random sample of clones (cDNA or genomic) is plated out, and duplicate nitrocellulose filters are made, each of which contains a representative imprint of all of the clones on the plate. Each filter is hybridized to a probe made from mRNA from one of the two cells to be compared. The probe may be made by fragmenting and then directly labelling the mRNA, or it may be made by reverse transcribing the intact mRNA into cDNA. Generally the cDNA has a much higher specific activity than the mRNA, and therefore makes a more effective probe.

Clones which give a signal with one of the probes, but not with the other are inferred to contain sequences which are differentially expressed in the corresponding cells. The plus-minus method of screening has one major limitation; only those clones which give a positive signal with one of the probes are chosen. The clones which contain sequences in the low abundance class may not give a signal with either probe, although they may still be differentially expressed. Various authors have estimated the abundance of the sequences which have been detected using plus-minus screening (see Table 1). The figures generally fall into the range of 0.1% (corresponding to the moderate abundance class of transcripts), although St. John and Davis (1979) found they could detect transcripts at ten-fold lower abundance (corresponding to the 'higher' low abundance transcripts). These authors used a high specific activity cDNA probe, and screened phage libraries, and both of these factors appear to increase the sensitivity of the screening. In contrast, screening plasmid libraries (colony hybridization rather than plaque hybridization) generally appears to increase the background and so decrease the sensitivity of the screening process.

3. *Hybridization Competition*

The plus-minus screening approach rests on the assumption that very similar quantities of DNA representative of each recombinant clone is present on each of the duplicate filters. Furthermore, the individual clones must be well isolated so that they can be identified unambiguously with the corresponding signal from each of the duplicate filters. This means that the libraries to be screened must be plated at a low density or in an ordered array, and that false positives can arise if unequal amounts of DNA are transferred to the duplicate filters. Several methods have been developed which require screening of only a single nitrocellulose copy of the clones of interest, and at a density higher than is possible with plus-minus screening. The simplest of these is RNA hybridization competition (Mangiarotti *et al.*, 1981). mRNA from the

cell of interest is labelled using [gamma^{32}P]-ATP and polynucleotide kinase. This is mixed with a large excess of total cytoplasmic RNA from a second cell displaying the biological differential under investigation, and the mixture used directly to probe the clones of interest. Sequences which are common to both cells will be labelled, but will be diluted by the excess of unlabelled competitor and give no signal; whereas sequences present only in the labelled mRNA will not be diluted and will detect clones containing homologous sequences. If clones are at a high density they can be isolated by replating at a lower density and rescreened using a similar preparation of probe. RNA hybridization competion can probably detect sequences in the moderate to 'high' low abundance class, with a sensitivity similar to the plus-minus method. However, the limits of detection using this method have probably not been determined, since high specific-activity cDNA probes could equally well be used and would likely increase the sensitivity (the competitor mRNA would now dilute out the target sequences for the cDNA probe but would have the same overall effect). Alternatively the labelled cDNA and competitor RNA could be prehybridized in a small volume to a high Rot value, so that sequences common to both would form hybrids and simply be unavailable to hybridize to clones.

4. *Subtractive Hybridization to Generate cDNA Probe*

Subtractive hybridization is a second method which can be used to prepare differential gene specific probes which can be used to screen a single nitrocellulose copy of the clones of interest, plated at a high density. This method is an obvious outgrowth of the methods described in the previous section. Labelled cDNA is prepared from the cells of interest, the template mRNA is removed by base hydrolysis, and the cDNA is hybridized with a large molar excess of mRNA from a second cell whose biological differential is of interest. The hybridization is done in a minimum volume, and allowed to continue to a high Rot value. Sequences which are expressed

by both cells will be capable of forming double stranded molecules, and the cDNA which remains single stranded can be separated by hydroxyapatite chromatography. The single stranded cDNA probe is thus highly enriched for sequences differentially expressed in the original cell type, and the subtractive hybridization can be repeated to gain further enrichment (cascade hybridization; Timberlake, 1980). The lowest abundance sequences which can be detected with subtracted probes are reported to be in the order of 0.001% (Hedrick *et al.* , 1984a), which would correspond to mRNAs of the lowest abundance class. The reason why this method of detecting differential genes is about ten-fold more sensitive than the plus-minus method is not immediately clear. One plausible theory is the following. The activity concentration (*i.e.* the ^{32}P counts of probe per ml of hybridization solution, given near maximal specific activity) of cDNA probe routinely used in plus-minus screening is usually too low to give a positive hybridization signal with sequences corresponding to the lowest abundance class of mRNAs. If the activity concentration of the probe was increased ≈20 fold, low abundance sequences might be detected, but the large amounts of probe used would also increase the background signal by a corresponding amount, thus obscuring the plus-minus effect. A subtracted probe, on the other hand, increases the activity concentration of the probe specific for differential genes >20 fold, while keeping constant or even lowering the total amount of radioactive probe used in the hybridization; thus the signal to background ratio is much higher. This idea suggests that hybridization competition might also achieve higher levels of sensitivity, if very high activity concentrations of probe are used.

5. *Subtractive Hybridization to Generate Subtracted Libraries*

The ability to isolate differential gene sequences physically by subtractive hybridization means that these sequences themselves can be cloned, to make a 'subtracted library' which is highly enriched for the sequences of interest. Subtractive

cloning of this kind was first done by Sargent and Dawid (1983), using a single cycle of hybridization and hydroxyapatite separation as described above. This subtracted cDNA library contained differential genes derived from the lowest abundance class of mRNAs. These results are not surprising, since if the differential sequences (gastrula vs. oocyte, see Table 1) represented (for example) >6% of the mRNAs, and since the library was enriched 15 fold for differential sequences, then such sequences would theoretically be present in >90% of the clones in the library. If the differential genes are distributed evenly into the three abundance classes of mRNAs, then roughly one-third of all the clones in the library will contain low abundance, differential sequences. Indeed the authors found that 84 randomly chosen clones all gave a higher or negative signal when probed with gastrula cDNA probe compared to oocyte cDNA probe (clones which give a negative signal represent low abundance gastrula-specific sequences). Davis *et al.* (1984) constructed a B-lymphocyte specific cDNA library by subtracting T-lymphocyte sequences. These cells differ by ~2% of their mRNAs, so that 40-80% of the clones in the subtracted library (estimated to be enriched 20-40 fold) contained B-cell specific sequences. Clearly subtractive cloning represents a powerful method for making a permanent collection of differential genes, which can then be studied exhaustively.

6. *Subtracted cDNA Probe Used on Subtracted Libraries*

If a subtractive library is not highly enriched for differential sequences, differential genes can still be isolated if the library is screened with probe also prepared by subtractive hybridization. This combined approach was used by Hedrick *et al.* (1984a) to isolate the gene encoding the ß-chain of the T-cell antigen receptor (Table 1). In this case the subtracted library was enriched ~20-fold for T-cell specific sequences (vs. B-cell), and thus ~40% of the clones represented T-cell specific sequences. The cDNA probe was made from a special subset of the T-cell mRNAs (those coding for

secreted or membrane-bound proteins), and was then depleted of sequences common to B-cells by subtractive hybridization. (In fact a number of T-cell specific clones were isolated by this method. The T-cell receptor gene was identified from among these by further demonstrating that the corresponding sequences were rearranged in the genome of T-cells compared to B-cells or liver cells.) Screening a subtracted library with a subtracted probe is not significantly different from screening an unselected library, but the subtracted library contains a higher proportion of differential genes and this means that fewer clones need to be screened. This can be a particular advantage since the amount of subtracted probe is often limiting, so that only a few filters can be screened.

7. *Cloning Low Abundance Differential Genes- Which Method is Appropriate?*

The difficulty of finding cDNA clones containing differential genes depends on a number of factors. If the task is simply to discover any differential gene which may be present, the magnitude of the task depends on the size of the differential. If the differential represents 25% of the mRNAs, clones can be picked at random and one-quarter will contain the desired gene. Depending on the abundance distribution of the differential genes, perhaps as many as half of the differential clones will contain low abundance sequences (*e.g.* brain-specific sequences). In this case the size of the biological differential is very large, however, so that the task of discovering the biological significance of any given differential sequences is difficult. If the differential is small (say 2% of the mRNAs), then a random strategy is unlikely to yield a differential gene, and another approach is required. The plus-minus strategy is reasonable, but the set of differential genes chosen will generally be limited to the high and moderate abundance class mRNAs. These may represent more than half of the total quantity of differentially expressed mRNA, and constitute a biologically interesting and important set of sequences. However, if the investigator wishes to discover low abundance differential genes, which may also be biologically significant and interesting,

then a subtractive strategy is apparently required. Subtracted probe will (in theory) identify the total set of differential genes, and a highly enriched subtracted library will (in theory) constitute the total set of differential genes. (This type of simplified analysis does not include differential genes from closely related gene families, which would not appear in a subtracted library.) If clones are chosen randomly from this set, the proportion which fall into the low abundance class will depend on the abundance distribution of the differential set. This may be very different from the abundance distribution of the set of mRNAs expressed by the 'typical mammalian cell' described above. For example, if a differential represents 4% of the total mRNAs, and includes one high abundance sequence, four moderated abundance sequences, and 70 low abundance sequences, then 84% of the differential clones picked would contain the same high abundance sequence (12,000 mRNAs of the total differential of 14,250 mRNAs are the same), 9% would contain moderate abundance sequences ([4 x 300]/14,250), and 7% would contain low abundance sequences ([70 x 15]/14,250). These very hypothetical calculations have a very practical conclusion. The size and abundance distribution of the differential set of genes will determine the method of choice and the likely outcome if one simply sets out to clone differential genes from the low abundance class.

8. *Other Methods for Cloning Differential Genes*

Table 1 also includes examples of a number of other strategies which have been employed to clone differential genes. Stetler and Thorner (1984) cloned a number of genes which are induced in yeast following exposure to mating pheromone. Cells were washed and cultured briefly in the absence of uridine, following which 4-thiouridine was added, and then the mating pheromone. Newly synthesized mRNAs, including many which were induced by the pheromone, incorporated thiouridine and could therefore be recovered from total mRNA by phenylmercury agarose chromatography.

RNA was also prepared in parallel from untreated yeast cultures, and two resulting cDNA probes were used to screen a genomic library in a plus-minus fashion.

Cloning of the T4 and T8 molecules (Table 1) was done by transfecting a mouse fibroblast cell line (L-cells) with human genomic DNA, staining the resulting transformants with antibody to the surface marker of interest, and then isolating transformants which scored positive. In this case a biological differential was artificially created by DNA-mediated gene transfer. The human genes of interest could not be recovered directly since they did not contain Alu sequences (a sequence which is characteristic to human DNA, and is repeated throughout the genome), and hence a subtractive approach was used.

Genetic variants that arise spontaneously also provide examples of biological differentials (in contrast to those which are artificially created, as described above). The mutant may contain only point mutations or very small deletions, but these may lead to the overproduction or underproduction of one or more mRNAs. The resulting differential genes can be cloned, and may often (but not necessarily) lead to the discovery of the original mutation. In any case the differential genes are of interest since they often contribute to the mutant phenotype. For example, Atkinson *et al.* (1985) discovered a sequence which was overproduced in a mutant strain of aspergillus using the plus-minus strategy.

The differential genes discussed above are identified as such because they are differentially expressed at the RNA level. Genes can also be identified which are differential at the genomic level. For example, male cells contain a Y chromosome not found in female cells, and the sequences which are unique to the Y chromosome can be cloned. An interesting strategy to do this was developed by Lamar and Palmer (1984), which they call deletion enrichment and cloning. Briefly, male DNA is cut with a restriction enzyme, melted, and mixed with a 100-fold excess of denatured female DNA

which has been sheared at random by sonication. The DNA is allowed to anneal to a high Cot, and is then ligated into a vector which has been cut to give 'sticky ends' compatible with those in the male DNA. Any male sequences which are homologous to female will hybridize to these and not result in DNA segments with 'sticky ends.' Any male sequences which are unique to the male will hybridize only with their original complement, and when they do so they will restore the 'sticky end.' Only DNA segments with 'sticky ends' will be ligated into the vector, and these sequences will appear as clones in the library. This strategy enabled Lamar and Palmer (1984) to isolate a number of genomic clones which contained sequences found specifically on the murine Y chromosome. The same strategy can be used to clone those sequences which are present in a normal individual, but which are absent in a mutant individual who has a significant chromosomal deletion (homozygous or heterozygous). Genes of this type were first cloned by Kunkel *et al.* (1985), who also used a phenol emulsion reassociation technique (PERT; Kohne *et al.* , 1977) to increase significantly the rate of DNA reassociation during the annealing step.

The first part of this chapter introduced the idea of a differential gene, and then went on to discuss how the set of mRNAs in a given cell can be characterized, and how the set of mRNAs expressed in two different cells can be compared. This process of comparison is necessary to define and characterize the set of differential genes. The present section describes a number of strategies and methods which can be used to clone differential genes. The final portion of this chapter will consider a specific biological system where the idea of differential genes, and methods for cloning differential genes can be applied.

D. DIFFERENTIATION OF LYMPHOCYTES DURING AN IMMUNE RESPONSE

1. *Cells of the Immune System and Immune Recognition*

The immune system consists of a complex network of cells which protect the host from infection by foreign micro-organisms. Three major cell types are involved in the immune response; these are macrophages, granulocytes, and lymphocytes. The lymphocytes have been divided into two classes, T-cells and B-cells, according to their site of maturation during the ontogeny of the immune system. T-cells and B-cells have different functions and consequently different surface molecules, and these can be used to distinguish the two types of lymphocytes.

'T-cells and B-cells are capable of specific antigen recognition in the vertebrate immune response. B-cells recognize free antigen through cell surface bound immunoglobin, with no other requirement except antigen/receptor complementarity. T-cells recognize antigen only when it is presented on the cell surface in the context of class I or class II molecules encoded by the major histocompatibility complex. This interaction is said to be self-MHC restricted in that only antigen presented on cells expressing self-MHC alleles can be recognized. Three functional classes of T-cells have been defined: T helper cells, which stimulate immune responses; T supressor cells which diminish immune responses; and T cytotoxic cells, which are involved in the direct killing of cells expressing antigenic molecules. T helper cells usually recognize antigen associated with class II MHC molecules expressed selectively on antigen presenting cells (APC) such as macrophages or B cells. T cytotoxic cells (CTLs) generally respond to antigens of cellular origin, such as viral or tumor antigens in association with the class I molecules present on most cells. The mode of recognition of T suppressor cells has not been well characterized' (Goverman *et al.*, 1986).

2. *Lymphocyte Activation*

The various cells of the immune system occur in organized tissues and organs (*e.g.* spleen, lymph nodes, Peyer's patches), and are also present as a recirculating pool in the blood and lymph. Normally the majority of the various types of lymphocytes are

in a resting state, and are not activated to cause an immune response. However, if specific lymphocytes encounter a foreign antigen which they can recognize (the recognition event is subject to the constraints outlined above), these cells become activated so that they now mount an immune response in an attempt to destroy the foreign substance (or cell). During the activation event, a specific set of new genes is expressed by the lymphocyte, and the products of these genes are responsible for the specialized immunological functions of the activated cells. If one compares the set of genes expressed by a particular lymphocyte before and after the activation event, the induced genes represent a set of differential genes similar to the genes discussed in the first part of this chapter. It is this set of differential genes, and in particular those expressed in activated T-helper cells, which will be considered in this thesis.

Following binding of a specific antigen to the surface immunoglobin present on a B-cell, the cell is activated to both proliferate and then mature, to produce cells which secrete free antibodies with the same binding specificity as the original surface immunoglobulin. In general, however, both the proliferation of the B-cell and its maturation into an antibody secreting cell depend on soluble factors released by T helper lymphocytes in proximity to the B-cells (Kishimoto, 1985). (The exception to this generalization is Lyb 5^+ B-cells [Paul, 1984].) Furthermore, as the antigen-secreting B-cell matures it may switch the class of immunoglobulins secreted (*e.g.* from IgM to IgG), and this class switch is also believed to be dependent on soluble (or cell surface) molecules produced by T-helper cells in the vicinity (Tesch *et al.*, 1986).

Cytotoxic T lymphocytes become activated after they bind to their target cells. This binding is mediated by the T-cell antigen receptor (reviewed in Kronenberg *et al.*, 1986), and is subject to the class I MHC restrictions described above. Very little is known about the genetic program or the lytic mechanisms that are activated once the cell recognizes its target. It has been suggested that the activated cytotoxic cell may destroy

its target by secreting a pore forming protein into the contact zone between the effector and target cells (Young and Cohn, 1986). Lobe *et al.* (1986), and Gershenfeld and Weissman (1986) have discovered CTL-specific cDNAs which code for novel proteases. The transcription of these protease genes appears to precede the o. . . of cytolytic activity, and it has been suggested they may play a role in CTL activation. A number of experiments have shown that, in addition to target binding, the activation of CTLs may require one or more soluble factors made by T helper cells (Wagner and Rollinghoff, 1978). A recent report also indicates that a T-helper cell derived soluble factor may activate CTLs independent of antigen binding (Milanese *et al.*, 1986).

The idea of activation of specific T-supressor cells from a resting state to one where they are actively suppressing one or more immune responses has been considered in the literature (Tada, 1984). However, in general such an activation event is poorly understood, and will probably continue to be so until specific suppressor factors have been purified and molecularly cloned.

The activation of T-helper cells requires the presence of an accessory cell known as an antigen presenting cell (reviewed by Schwartz, 1984). Antigen presenting cells are cells which express MHC class II molecules on their surface, and include macrophages, B-cells, epithelial Langerhans cells and the specialized epithelial cells of the thymus. It is now apparent that these cells function by breaking down and denaturing large antigen molecules ('antigen processing'), and then affixing the resulting peptides to the class II molecules on their surface (Germain, 1986). The T-helper cell is activated when it binds to the combination of the class II molecule and the processed antigen. This binding is mediated by the heterodimer T-cell antigen receptor (Kronenberg *et al.*, 1986; Goverman *et al.*, 1986), in association with three invariant polypeptide chains identified by the T3 monoclonal antibody (Van der Elsen *et al.*, 1984). Activation of the T-helper cell also requires a soluble macrophage product,

interleukin 1. Figure 1 gives a schematic representation of the process which occurs during the activation of a T-helper cell. The figure also indicates that the requirement for processed antigen plus class II molecule binding to the T-cell antigen receptor can be bypassed *in vitro* by the addition of mitogens (such as concanavalin A or phytohemagglutinin), or by anti-T3 antibody which binds to the T3 complex associated with the T-cell antigen receptor. Furthermore, the requirement for interleukin 1 can be bypassed *in vitro* by the addition of the tumor promoting phorbol ester PMA (phorbol-12-myristate-13-acetate).

The activation of T-helper cells causes the cells to proliferate and to secrete a number of soluble factors (lymphokines) which in turn act on the various cells of the immune system to effect an immune response. For example, lymphokines produced by T-helper cells are known to be required for the growth and differentiation of antibody secreting B-cells. Similarily, T-helper derived lymphokines are apparently required for the growth and activation of cytotoxic T-cells. In this way T-helper cells have primarily a regulatory function, rather than a direct effector function during an immune response. A number of the lymphokines produced by activated T-helper cells are shown on Figure 1, and some of these will be discussed briefly below.

It has recently been shown that not all activated T-helper cells produce the same set of lymphokines, and that they can be classified into two major groups according to the lymphokines they produce (Mosmann, 1986). For the purposes of this discussion, however, T-helper cells will be considered as a single group. In general, when T helper cells are activated, they begin to express a new set of genes, and this induced set of genes includes a number which code for lymphokines.

3. *Lymphokines Secreted by Activated T-Helper Lymphocytes*

Perhaps the best known of the lymphokines secreted by activated T-helper cells is interleukin 2 (IL2). This lymphokine has been well characterized, both biochemically

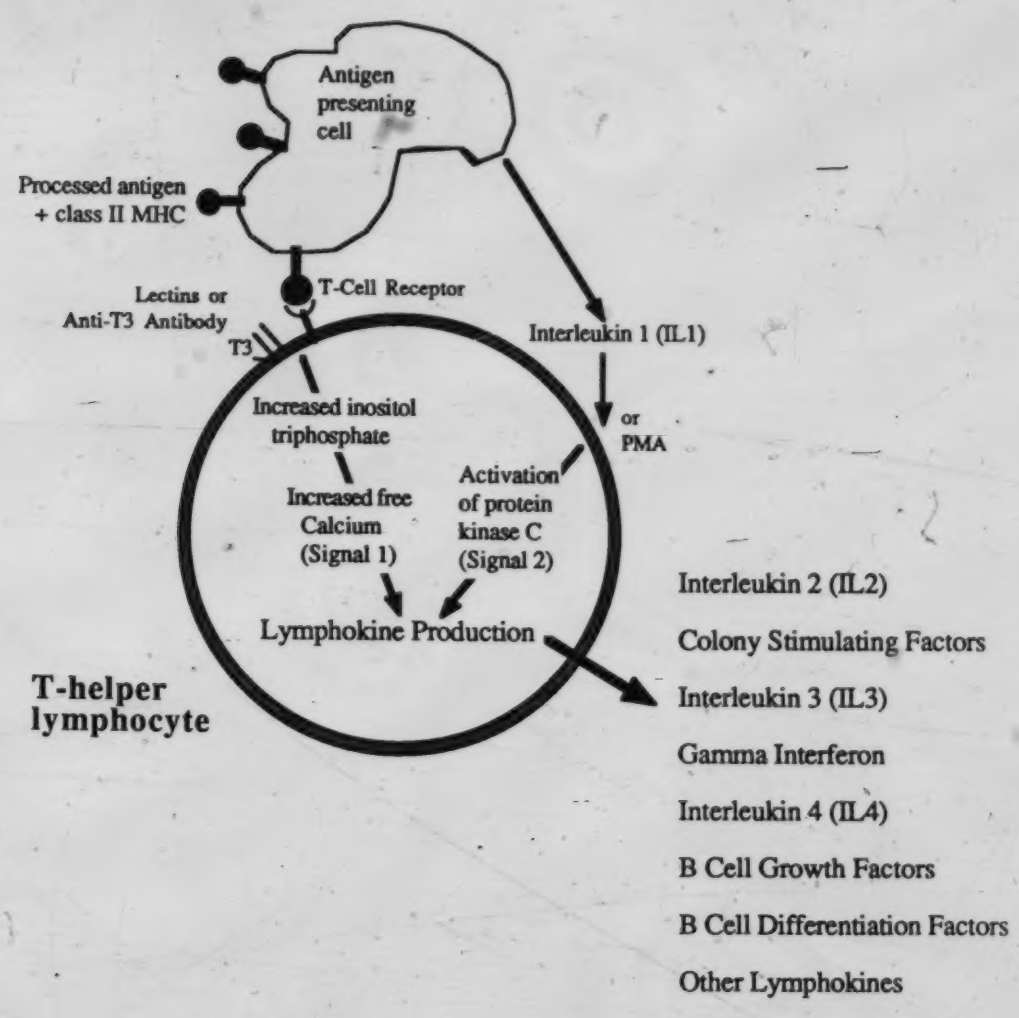


FIGURE 1. The activation of T-helper lymphocytes results in the production of a number of lymphokine molecules. The activation process is believed to require two signals, as illustrated above. When the T-cell antigen receptor-T3 complex binds to processed antigen in combination with self class II MHC molecules, a series of events results in the mobilization of intracellular calcium. The antigen requirement can be bypassed by lectins such as Con A, or by anti-T3 antibodies. The second signal occurs when interleukin 1 binds to receptors on the T-cell surface, and results in the activation of protein kinase C. The requirement for interleukin 1 can be bypassed by the phorbol ester PMA, which is known to activate protein kinase C directly. In normal helper lymphocytes these two signals in concert result in the secretion of a number of lymphokine molecules, and not all of these have been well characterized. The ideas about lymphocyte activation upon which this figure is based are found in Truneh et. al., 1985; Weiss et al., 1984a; Weiss et al., 1984b; Nishizuka, 1984; and Weiss et. al., 1986.

and in terms of its immunological function (reviewed in Farrar *et al.*, 1982). The cDNA's coding for both human and mouse IL2 have been cloned (Taniguichi *et al.*, 1983; Kashima *et al.*, 1985), and the recombinant proteins have been produced using both procaryotic and eucaryotic expression systems. The human cDNA sequence codes for a protein of 153 amino acids (with a signal sequence of 20 amino acids), and mouse IL2 protein is 169 amino acids in length. IL2 appears to have biological effects on many different populations of lymphocytes, but its primary effect is in promoting the growth of responding T-cells. To exert its biological effect, IL2 must interact with specific high affinity membrane receptors (Greene and Leonard, 1986) which are expressed on the surface of the relevant cells following activation by antigen or mitogen (Leonard *et al.*, 1985). IL2 causes the *in vitro* proliferation of cytotoxic T-cell lines, and this is the basis for a number of biological assays for the lymphokine (Gillis *et al.*, 1978); IL2 is known to induce proliferation of mitogen treated thymocytes and to augment the proliferation and generation of cytotoxic cells by alloantigen-stimulated T-cell populations (Farrar *et al.*, 1982);. IL2 also augments the growth of helper T-cell lines in the presence of syngeneic antigen presenting cells and antigen (Fathman and Fitch, 1984). A large number of other biological effects have been attributed to IL2, such as an indirect effect on B-cell growth and differentiation (Inaba *et al.*, 1983; Howard *et al.*, 1983), and an effect on the induction of natural killer cell and lymphocyte-activated killer cell activities (Henney *et al.*, 1981). This latter effect suggests that IL2 may have a use in the immunotherapy of cancer, and clinical trials of this form of therapy are currently underway (Rosenberg, *et al.*, 1986).

Activated T helper cells also secrete a number of glycoproteins which cause isolated bone marrow precursor cells to form colonies *in vitro* . The best known of these colony stimulating factors (CSFs) is GM-CSF, which stimulates the growth of granulocyte and macrophage colonies (Metcalfe, 1985). The cDNAs for murine and

human GM-CSF have been cloned (Gough *et al.*, 1984; Wong *et al.*, 1985; Lee *et al.*, 1985); the (mature) murine protein is ≈124 amino acids, and the (mature) human protein ≈127 amino acids in length. Another colony stimulating factor known as multi-CSF causes the proliferation of a wide range of cells, including essentially all of the cell types found in the bone marrow. Interleukin 3 (IL3) is of T-cell origin, and although included in the multi-CSF complex, it does not act on erythroid or stem cell populations (Metcalf, 1986). The cDNA for murine IL3 has been cloned (Fung *et al.*, 1984; Yokota *et al.*, 1984), and the processed polypeptide apparently contains 140 amino acids.

Activated T-helper cells also produce a number of other lymphokines (Figure 1). Immune interferon (gamma interferon; murine) consists of a peptide of 136 amino acids, the sequence of which was derived from the cDNA sequence (Gray and Goeddel, 1983). Gamma interferon stimulates the phagocytic and tumoricidal activities of macrophages, and increases the expression of MHC class I and II molecules on macrophages (Steeg *et al.*, 1982; Nakamura *et al.*, 1984; Paulnock-King *et al.*, 1985). Several lymphokines made by activated T-helper cells have an effect on the growth and differentiation of B-cells, and these are known as B-cell growth factors (BCGFs) and B-cell differentiation factors (BCDFs) (Howard and Paul, 1982; Kishimoto, 1985). The nomenclature in this field is still evolving, and can be confusing. For example, BCGF-1 is required for the proliferation of B-cells following the binding of antigen or anti-IgM antibodies to their surface immunoglobulins, but this lymphokine alone will also stimulate resting B-cells, and it has therefore also been called B-cell stimulatory factor (BSF-1). The cDNA for this protein (murine) has recently been cloned (Lee *et al.*, 1986; Noma *et al.*, 1986), and the recombinant protein has been shown to have a large number of effects. These include the promotion of growth in anti-IgM activated B-cells, differentiation in certain B-cells (to IgG1 and IgE secretion), expression of

class II MHC molecules in B-cells and growth of mast cells and certain T-cells. This lymphokine has been named interleukin 4 (IL4), and the cDNA sequence demonstrates that the precursor polypeptide is 140 amino acids in length.

4. *Isolating New Lymphokines by cDNA Cloning*

With both IL2 and IL4 it is clear that a single lymphokine can have a large number of biological effects, and furthermore, that a single biological effect (*e.g.* promoting T-cell growth) can be caused by more than one lymphokine. This makes the purification of these molecules a critical step if we are to understand how the various lymphokines participate in and regulate the immune response. This process is all the more difficult considering that very small quantities of lymphokines are secreted by normal lymphocytes. One strategy to isolate and purify the various lymphokine molecules would be to clone the cDNA molecules coding for them first, and then use recombinant DNA technology to express the proteins (using either procaryotic or eucaryotic expression systems). Activated T-cells make a large number of 'other lymphokines' (Figure 1) which have not yet been discovered, or if discovered, are in the most preliminary stages of characterization. This makes a direct cloning approach to the study of lymphokines even more attractive. The problem is how to isolate cDNA sequences which code for lymphokines, when nothing is known about the proteins. This thesis suggests that one solution to this problem may be to use a differential cloning approach. Resting T-helper cells do not produce lymphokines, whereas activated T-helper cells do; the genes which are induced during activation represent a differential set which can be cloned directly and should be highly enriched for cDNA's coding for lymphokines. Moreover, the cloning of non-lymphokine genes which are induced during the activation of T-helper cells may also be useful in understanding the biology of T-helper cells and the process by which they become activated.

5. *Tumor Models for Helper T-cell Activation*

Purified normal T helper lymphocytes are difficult to obtain in large numbers, and when obtained still represent a very heterogenous population of cells. For the purposes of the work described in this thesis large numbers of cells, preferably of clonal origin, were required. For this reason it was decided to use a tumor cell line which can serve as a model for helper T-cell activation. This is the murine thymoma cell line EL4 (Farrar *et al.*, 1980). Ordinarily EL4 cells grow rapidly in culture (doubling time ≈12 hours) in non-adherent grape-like clusters, and secrete no known lymphokines. However, following addition of the tumor promoting agent PMA, the cells stop growing, become adherent and more granular in appearance, and begin to secrete a number of lymphokines including IL2, GM-CSF, IL3, BCGF-1, BCDFμ, and a number of other factors which have not been well characterized (Paetkau *et al.*, 1984). Thus PMA treated EL4 cells appear to mimic T-helper cells which have been activated by antigen and IL1 (Figure 1), but EL4 cells appear to require only one of the two signals (PMA activates protein kinase C) which are normally required (Fujita *et al.*, 1986). Despite the dramatic induction of specific lymphokine expression, the overall expressed repertoire of EL4 cells is not altered greatly by stimulation with PMA (Paetkau *et al.*, 1984). A second human T-leukemia cell line (Jurkat) can be induced to secrete lymphokines by the addition of mitogens and PMA (Gillis and Watson, 1980), and this cell line was also used in some of the preliminary studies of lymphokine gene expression described in Chapter III.

E. THESIS OBJECTIVE

The objective of this study was to identify and characterize a number of the mRNAs which are induced in the mouse T-lymphoma cell line EL4 when it is treated with PMA. To achieve the objective, recombinant DNA techniques, together with

subtractive hybridization were used to isolate cDNA clones representative of the PMA-induced mRNAs.

CHAPTER II
MATERIAL AND METHODS

A. CELL CULTURE

1. *Cell Culture Medium*

The tissue culture medium referred to as RH is based on RPMI 1640 (Gibco Laboratories) with the addition of 20 mM sodium bicarbonate, 0.34 mM sodium pyruvate, 0.02 M HEPES (N-2-hydroxyethylpiperazine-N'-2-ethansulfonic acid) (Sigma) adjusted to pH 7.3, and supplemented with 100 I.U./ml penicillin G potassium (Ayerst) and 100 µg/ml streptomycin sulfate (Gibco). RH with 10^{-4} M ß-mercaptoethanol (Baker) is referred to as RHM, and RHM with 10% (v/v) heat inactivated fetal calf serum (Gibco) is referred to as RHFM. All cell cultures were incubated under 5% CO_2 in air at 37°C, 100% humidity unless otherwise stated. Large cultures were grown in standard glass T-flasks, and all other tissue culture was done in disposable plastic flasks (COSTAR, Nunc, or Corning) as appropriate.

2. *Cell Lines*

EL4.E1 is a subclone of the murine T-cell lymphoma described by Farrar *et al.* (1980) which secretes high levels of IL2 when cultured in the presence of tumor promoting phorbol esters. JEH.3B is a subclone of the human T-cell leukemia line Jurkat (Gillis and Watson, 1980) which was selected for its ability to secrete high levels of IL2 upon stimulation with T-cell mitogens (concanavalin A or phytohemagglutinin). This human T-cell line gives maximal IL2 production when mitogens plus phorbol esters are added to the culture. The COS-7 cell line is a simian fibroblast cell line which has been transformed by an origin-defective mutant of SV40 (Gluzman, 1981), and was obtained from ATCC (Bethesda). The L1210 cell line is derived from a chemically induced B-cell leukemia in mouse strain DBA/2 (Michalides *et al.*, 1982), and was

obtained from Bill Pohajdak. The cells used for the TCGF assay were an IL2 dependent CTL line, MTL2.8.2, of CBA/J origin, which were derived by immunization *in vivo* and, repeatedly, *in vitro* with BALB/c x CBA/J spleen cells (Bleackley *et al.*, 1982). The COS 7 and MTL2.8.2 cells grow adherently and were passaged accordingly using 25mM EDTA (MTL), or EDTA and trypsin (COS 7) to loosen monolayers. All cell lines were maintained in culture in RHFM, seeding at a density of 1-2 x 10^4 cells/ml, and harvesting at a maximum density of 2 x 10^6 cells/ml. Cell lines were also frozen at a density of 1 x 10^6 cells/ml in RHFM containing 10% dimethylsulfoxide (DMSO) (Fisher Scientific) and 20% fetal calf serum in 2 ml pro-vials (Cooke Laboratories, Alexandria, Virginia), by placing vials at -70°C overnight in a styrofoam box and then transferring them to liquid nitrogen for long term storage. Periodically (every two to three months) ongoing cell cultures were replaced with frozen stocks. These were thawed at room temperature, diluted into 10 ml RHFM, centrifuged at 250 g for 7 minutes and resuspended in fresh media.

3. *Reagents*

HPLC purified samples of cyclosporin A (CsA) were obtained from Dr. F. Pasutto (Department of Pharmacy, University of Alberta). The cyclosporin was dissolved in DMSO at a concentration of 5 or 10 mg/ml and when required was diluted to desired concentrations in RHFM. Phorbol 12-myristate 13-acetate (PMA) (Sigma) was dissolved in ethanol (1 mg/ml) and diluted with RH to a 5 µg/ml stock. Concanavalin A (Con A) (Calbiochem) was made up in RH at a concentration of 2mg/ml and filter-sterilized. Cycloheximide (Sigma) was made up in phosphate-buffered saline (PBS) at a concentration of 500 µg/ml. All these reagents were stored at -20°C. DEAE-dextran (Sigma #9885; 5,000,000 mw, chloride form) was dissolved in RH at a concentration of 2 mg/ml and filter sterilized. Dexamethasone (K-line pharmaceuticals, Downsview, Ontario) was diluted in RHFM to a concentration

of 500 μg/ml. These last two reagents were stored at 4°C.

4. *Stimulation of Cell Lines to Produce Lymphokines*

In order to induce IL2 synthesis, healthy EL4.E1 cells in log-phase growth (density not greater that 1 x 10^6 cells/ml) were resuspended at a concentration of 1 x 10^6 cells/ml in fresh RHFM containing 20 ng/ml PMA. To induce maximal IL2 production by the Jurkat cell line, healthy cells in log-phase growth were resuspended at 1 x 10^6 cells/ml in RHFM containing both PMA (20 ng/ml) and Con A (30 μg/ml). Cell pellets (for RNA extraction) and supernatants (for IL2 bioassay) were separated by centrifugation at 250 x g for 12-15 minutes. Cells were harvested at various times, generally between 12-24 hours, as will be indicated.

5. *Biological Assay for IL2 Activity*

The standard assay for IL2 is based on the ability of a test sample to induce proliferation of the IL2-dependent cell line MTL2.8.2. This cell line will grow only if IL2 is present, and rapidly dies (24-48 hours) if the lymphokine is absent. Two methods of measuring cell proliferation have been used; the first involves measuring the incorporation of [^{125}I]-dIUrd into newly synthesized DNA, and the second measures the ability of metabolically active cells to reduce the tetrazolium dye 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyl tetrazolium bromide (MTT) to a blue formazon product (Mosmann, 1983). Two-fold serial dilutions (in triplicate) of cell-free supernatants to be assayed were prepared in 96-well, flat-bottomed microtiter plates (Nunc), so that 50 μl of RHFM plus supernatant remains per well following the dilutions. 1 x 10^4 MTL2.8.2 cells in 50 μl of RHFM were then added per well, and the cultures incubated for either 24 hours (^{125}I-dIUrd assay) or 48 hours (MTT assay). For the MTT assay, four to six hours before measurement, 10 μl of a 5 mg/ml aqueous solution of the MTT dye (Sigma) was added per well. At the time of harvest, 20 μl of 10% N-lauroylsarcosine (Sarkosyl) (Sigma) was added to each well, and the trays were

incubated for an additional 30 minutes at 37°C in order to solubilize the cells. To dissolve the dye and bleach the phenol red color of the cultures, 80 µl of 0.04N HCl in 2-propanol (1:290 v/v) was added, and the samples were mixed for 2 minutes on a vibrating platform (IKA Schuttler MTS4, Janke and Kunkel). The color was allowed to develope for 15 minutes at room temperature, and the absorbance of each of the wells was determined and recorded automatically using a Dynatech MR600 ELISA plate reader (570 nm test and 630 nm reference wavelengths). For the dIUrd assay, 50 µl RHFM containing [5-^{125}I]-deoxyuridine (2 µCi/ml, 700 Ci/mM), 5-fluorodeoxyuridine (50 µg/ml), and inosine (0.1 mM) was added to each well 6 hours before harvest. Fifteen minutes prior to harvest the media was removed using a pasteur pipet under low vacuum, and 50 µl 20 mM EDTA was added to each well. Cells were collected on glass filters using an automated cell harvester (Titertek; Flow Laboratories), and the radioactivity measured. IL2 activity was expressed in terms of ED_{30} units, in which 1 ED_{30} U/ml generated 30% of the maximum response obtained with saturating IL2. This is the same as the dilution of IL2 that gave 30% of maximal response. (Recently IL2 activities in Dr. Paetkau's laboratory have been based on a laboratory standard, and are referred to as 'standard units' [Hooton *et al.*, 1985].)

B. ISOLATION AND CHARACTERIZATION OF RNA

1. *Large Scale Isolation of Total Cellular and Poly A$^+$ RNA*

After 16 hours of culture total cellular RNA was recovered from various groups of cells (2-3x10^9 cells per group) by homogenizing the cell pellets in 10-16 ml of 4 M guanidinium thiocyanate (BRL), 0.5% sodium N-lauroylsarcosine, 25 mM sodium citrate (pH 7.0), 0.1 M 2-mercaptoethanol, and 0.1% Sigma antifoam A (Chirgwin *et al.*, 1979). Homogenization was done in a Sorvall single-blade homogenizer, using three 30 second cycles with constant cooling on ice. 3.2 ml of the homogenate was

layered over 2 ml of 5.7 M CsCl, 0.1 M EDTA pH 7.0 and the RNA was pelleted through the cesium chloride by centrifugation in a Beckman SW 50.1 swinging bucket rotor (36,000 RPM x 12 hours at 20°C). The supernatant was aspirated and the pellets rinsed quickly and gently with ice cold, autoclaved water. The RNA pellets were dissolved in 10 mls of 25mM EDTA, 0.1% SDS with gentle heating (42°C), shaking, and vortexing, and debris was removed from the RNA solution by slow speed centrifugation (4000 RPM x 10 min. in Sorvall SS34). The RNA was made 0.3M in sodium acetate (pH 5.3) and precipitated by adding 2.5 volumes of 95% ethanol and storing at -70°C overnight. The pellet was collected by centrifugation (12,000 RPM x 30 min. in Sorvall SS34), dried in the lyophilizer, and dissolved in a small volume of 10 mM Tris- HCl pH 7.4, 1 mM EDTA, 0.1% SDS. LiCl was added to a final concentration of 0.5 M, and Poly A^+ RNA was isolated from this material by three consecutive cycles of adsorption to and elution from oligo-dT cellulose (Collaborative Research), using 0.5 M LiCl for loading, 0.1 M LiCl for washing, and LiCl free buffer for eluting the column, as described by Bleackley *et al.* (1981). The poly A^+ fraction was precipitated twice (using sodium acetate and ethanol) and finally resuspended in water. This material was used directly for making northern blots, for translation in wheat germ lysates or oocytes, for generating cDNA, for subtractive hybridization, and for cap labelling as described below.

2. *Small Scale Isolation of Cytoplasmic RNA and Poly A^+ RNA for Northern Analysis*

Total cytoplasmic RNA was harvested from 0.5-2x10^7 cells using this small scale isolation procedure. Cell pellets were washed once in PBS, pelleted, and resuspended in 90 µl of ice-cold 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (TE). Cells were lysed by the addition of two 10 µl aliquots of 5% Nonidet NP-40 (Sigma) with 5 minutes' incubation on ice between the additions. Nuclei were pelleted (15,000x g, 2

min in a microfuge), and the supernatant was transferred to a 1.5 ml tube containing 300 µl TE, 20 µl 10% (w/v) SDS, and 10 µl of 2 mg/ml proteinase K (BRL). This mixture was incubated for 30 minutes at 37°C, and RNA was precipitated with sodium acetate (0.3M, pH 5.3) and 2.5 volumes of ethanol (-70°C overnight). RNA was recovered by centrifugation in a microfuge and the pellet dried and dissolved in 25 mM EDTA, 0.1% SDS. This material was either run directly on gels to make northern blots, or used to obtain poly A^+ material. For the latter, 20µg of RNA (concentration determined by fluorimetry) was hybridized to 0.5 cm^2 of Hybond^{tm}-mAP paper (Amersham), and poly A^+ RNA extracted following the manufacturer's directions. Briefly, aliquots of the RNA solution were placed on the mAP pieces which had been soaked previously in 2X SSPE (0.15 M NaCl, 0.01 M monosodium phosphate, 1 mM EDTA, Maniatis *et al* ., 1982). The RNA solution was drawn through the mAP by placing the mAP pieces on several layers of Whatman 3 mm filter paper. The mAP pieces were washed separately in 0.5 M NaCl using a Costar 6-well dish (three 5 ml washes, each lasting 10 min at room temperature), and finally in 5 ml of 70% ethanol for 2 min. The air-dried mAP pieces were transferred into 1 ml of distilled water in a 1.5 ml tube, and heated to 70°C for 5 minutes to release the poly A^+ RNA. Yeast tRNA (20µg) was added to the tube as carrier, and the RNA was precipitated with ethanol and 0.3M sodium acetate as before. This material was redissolved in 25 mM EDTA, 0.1% SDS, and run on northern gels as described below.

3. *Translation of Poly A^+ RNA in Wheat Germ Extracts*

Wheat germ extracts treated with micrococcal nuclease were purchased from BRL, and used essentially as described (Harnish *et al*., 1986). Standard incubations (30µl) contained 50 mM KOAc, 0.7 mM $Mg(OAc)_2$, 50 µCi [^{35}S]-methionine (900-1100 Ci/mmol, NEN), 20 mM Hepes (pH 7.5), 1.2 mM ATP, 0.1 mM GTP, 5.5 mM creatine phosphate, 0.2 mg/ml creatine kinase, 80 µM spermidine, 50 µM each of

19 amino acids, 33% (v/v) wheat germ extract, and 1-6 µg of RNA. Translation mixes were incubated for 60 min at 25°C, and triplicate 8 µl aliquots were then assayed for IL2 activity.

4. *Translation of Poly A$^+$ RNA in Xenopus laevis Oocytes*

Female *Xenopus laevis* were obtained from NASCO Int. Corp. (Fort Atkinson, WI) and maintained in the Department of Zoology, University of Alberta. Frogs were sacrificed and the ovaries removed and placed in Modified Barths Solution (MBS) (88 mM NaCl, 1.0 mM KCl, 2.4 mM $NaHCO_3$, 0.82 mM $MgSO_4$, 7.5 mM Tris-HCl pH 7.6, 0.33 mM $Ca(NO_3)_2$, and 0.41 mM $CaCl_2$). Individual oocytes were 'teased' from the ovaries using blunt and sharp dissection. These oocytes were washed in MBS, and micro-injected with ≈50 nl of poly A$^+$ RNA (2 mg/ml in water), using a 12X dissecting microscope and a micromanipulator. Micro-needles were hand-pulled from custom glass pipets (N51A glass, O.D. 0.0665", I.D. 0.0555", Drummond Scientific, Broomall, PA), and injections were done under fluorescent illumination to prevent the needle from drying out. Following microinjection, oocytes were incubated individually in micro-wells (terizaki plates, Falcon Plastics) at room temperature in MBS (≈30 µl/oocyte), and harvested at 60 to 72 hours. The liquid surrounding the oocyte was harvested before removing the oocytes, and supernatants and oocyte lysates were assayed separately for IL2 activity. These numbers were combined to yield the total amount of IL2 produced per oocyte. In bioassays of translated proteins (wheat germ or oocyte), phenylmethylsulphonylfluoride (PMSF) at a concentration of 25 µg/ml was added to the RHFM to inhibit proteolysis of the translated products.

5. *RNA Gels and Northern Transfers*

RNA was separated on denaturing gels and blotted directly onto nitrocellulose or nylon membranes using established methods (Thomas, 1980). The running buffer contained 20 mM 3-[N-morpholino] propane sulfonic acid pH 7.0 (MOPS, Sigma), 5

mM sodium acetate, and 1 mM EDTA; the buffer was recirculated with a pump during electrophoresis (Maniatis *et al.*, 1982). Agarose gels (0.8-1.0%) were made up in this buffer and also contained 0.67% formaldehyde, and 1.0 µg/ml ethidium bromide. RNA samples (1-40 µg) were denatured by incubating at 55°C for 15 minutes in 20 µl of running buffer containing 6.5% formaldehyde and 50% deionized formamide (Fluka, Germany). The samples were chilled on ice, and 2 µl of 50% glycerol, 1 mM EDTA (with no dyes) was added to each sample prior to loading. Progress during electrophoresis was monitored by adding loading buffer containing dye (0.25% bromophenol blue) to one of the tracks which did not contain RNA, and gels were run until this dye had moved two-thirds of the length of the gel. Gels were destained 30-60 minutes in 20X SSC (Maniatis *et al* ., 1982), photographed, and transferred to either BA85 nitrocellulose filters (Schleicher and Schuell) or nylon membranes (Hybond-N, Amersham). Prior to transfer the membranes were soaked in 10X SSC, and blotting was done according to the method outlined in Maniatis *et al.* (1982). Following transfer the lane origins were marked on the filters and they were baked for 2 hours in a vacuum oven (80°C), and then stored at room temperature.

C. NUCLEIC ACID PROBES AND HYBRIDIZATION CONDITIONS

1. *Oligonucleotide Probes*

Oligonucleotide probes were synthesized by P. Barr (Chiron, Emeryville CA) on an Applied Biosystems nucleotide synthesizer. They were based on the published sequences for human (Taniguchi *et al.*, 1983) and mouse (Fuse *et al.*, 1984) IL2 and murine GM-CSF (Gough *et al.*, 1983), and their characteristics and designations are given in Table 5. These synthetic oligonucleotides were labelled at the 5' end using T4 polynucleotide kinase (16 U per reaction, Pharmacia) in an 80 µl reaction containing 0.5-1.0 µg of oligonucleotide, 150 µCi of [gamma-^{32}P]-ATP (New England Nuclear,

>3000 Ci/nmol), 50mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 0.1 mM spermidine, and 0.1 mM EDTA. The reaction was incubated at 37°C for 1 hour, and labelled DNA was separated from unincorprated ^{32}P by chromatography over Sephadex G-50F using disposable ~1.8ml columns (made from a small pasteur pipet plugged with glass wool). The buffer for these columns was 0.1 M NaCl, 50 mM Tris-HCl pH 8.0, 1mM EDTA, and 0.02% SDS.

2. *Labelling of cDNA Inserts and Other DNA Fragments*

DNA fragments which had been purified were labelled by nick translation (Rigby *et al.*, 1977) using a kit from Bethesda Research Laboratories (BRL, Gaithesburg, Maryland) and [alpha-^{32}P]-dCTP (New England Nuclear, >3000 Ci/mmol). cDNA inserts which were cut out of 1% low melting point agarose (BRL) gels were labelled directly (*i.e.* without further purification) using the 'random primer method' of Feinberg and Vogelstein (1983). Excised bands were placed in an Eppendorf tube, and an approximately equal volume of water was added to the tube. The tube was boiled for 7 minutes and then cooled to 37°C. 33 µl of this mixture was added to 10 µl of OLB, 2 µl of bovine serum albumin (BSA) (10 mg/ml, BRL), 50-100 µCi of [alpha-^{32}P]-dCTP (New England Nuclear, >3000 Ci/mmol), and 5 units of Klenow (large fragment of DNA polymerase I, Promega Biotec), and the reaction was incubated for 3-7 hours at 37°C. The OLB mixture consisted of 250 mM Tris-HCl pH 8.0, 25 mM $MgCl_2$, 45 mM 2-mercaptoethanol, 1.0 M Hepes buffer pH 6.6, 100 µM dATP, 100 µM dGTP, 100 µM dTTP, and random hexadeoxyribonucleotides (Pharmacia #2166) at 540 µg/ml. With either method of labelling, labelled DNA was separated from unincorporated base using exclusion chromatography as described above for the oligonucleotides. Final specific activity of all probes was about 1 x 10^8 cpm/µg.

3. *Prehybridization, Hybridization, and Washing Conditions for Blots*

Prior to prehybridization all filters were washed gently in 4X SSPE (Maniatis *et al.*, 1982). All filters were prehybridized at 42°C for 6-12 hours in a solution containing 50% formamide (v/v), 5X SSPE, 0.1% SDS, 1mM ATP, 5X Denharts solution, 200 µg/ml sheared salmon sperm DNA, 200 µg/ml yeast tRNA, and 50-100 µg/ml poly A (Sigma). For oligonucleotide probes, filters were then washed four times in 4X SSPE, and hybridized at 42°C overnight in a solution identical to the pre-hybridization mixture except it contained 20% formamide and 10% dextran sulfate (Pharmacia). For longer DNA probes, filters were hybridized at 42°C overnight in a solution identical to the pre-hybridization mixture except that it contained 1X Denharts solution and 10% dextran sulfate. 1X Denharts solution contains 0.02% w/v each of ficoll, polyvinylpyrollidone, and bovine serum albumin. The concentration of ^{32}P-labelled probes was 0.5 - 4.0 x 10^6 cpm/ml in all hybridizations. Filters which had been hybridized with synthetic oligonucleotide probes were washed three times (15 minutes each time) at 42°C in 2X SSPE (or 2X SSC), 0.1% (w/v) SDS. Filters hybridized to longer probes were washed at 55°C, three times in 2X SSPE (or 2X SSC), 0.1% SDS, and three times in 0.2X SSPE (or 0.2X SSC), 0.1% SDS, with 10 minutes for each wash. Filters were air dried, wrapped in Saran wrap, and exposed to film (Kodak X-Omat AR) with or without intensifying screens as necessary.

D. PREPARATION OF cDNA LIBRARIES IN PHAGE LAMBDA GT10

The method used to generate cDNA described below is based on the unpublished protocol of T. St.John, J. Rosen, and H. Gershenfeld (Department of Pathology, Stanford University), and includes many improvements and adaptations developed by D. Denny.

1. *Preparation of Cloning Reagents*

A number of reagents could not be purchased when I constructed my cDNA

libraries. The preparation of these reagents is described in this section.

a) Vanadyl Sulfate-nucleoside Complexes (Vnx). This reagent, which is a useful RNase inhibitor in both reverse transcription and translation reactions, was prepared by a method adapted from Berger and Birkenmeir (1979). A 2M solution of vanadyl sulfate was prepared by adding 2.6 gm of $VOSO_4$ (Fisher # V-8) to 8.0 ml of water, and dissolving by boiling in a 15ml corex tube. A 0.25 M adenosine solution was prepared by adding 3.34 gm adenosine (Sigma #A-9251) to 48 ml water, and dissolving by boiling. While still hot, the adenosine was placed in a 100 ml beaker (with stir bar) on a pH meter, and 6 ml of the hot 2M $VOSO_4$ solution was added dropwise with stirring. The formation of the vanadyl-adenosine complex was signaled by a sudden change in the color of the solution to dark blue, and a rapid drop in the pH to ≈2.5. This solution was immediately titrated to pH ≈6 by quickly adding a pasteur pipet full of 5N NaOH, followed by dropwise addition of the base. The volume of the solution was brought up to 60 ml by addition of 6 ml of water, and the pH titrated to 7.5 with 1N NaOH. The resulting 0.2 M stock solution was stored in 1 ml aliquots at -70°C and was stable for several months. Frozen aliquots were used immediately after thawing, and then discarded.

b) Phage Lambda gt 10 Arms. All bacterial culture media and plates were made up according to Maniatis *et al.* (1982). For growth of phage, standard LB was supplemented with 10 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, and 0.25% glucose. A stock of lambda gt 10 was obtained from D. Denny (Department of Microbiology, Stanford University), and was titred on the bacterial strain KM392. Twenty eight individual cloudy plaques were picked from a plate of suitable density, and each was placed in a separate tube containing 1 ml of Ø-dil (10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$). The phage particles were allowed to elute at room temperature for several hours, then 1ml of fresh log phase KM392 cells ('plating stock,' grown in 0.2%

maltose, resuspended in Ø-dil at an $O.D._{600}$ of 2.4) were added to each tube. The phage particles were allowed to attach at room temperature for 30 minutes, then each tube was plated onto a single large (150 mm) plate, using 6 ml of 'top agar' per plate. The plates were incubated inverted at 37°C for 7 hours, at which time they were confluent. Any clear plaques present on the lawns were removed with a sterile pasteur pipet, and a plate with a large number of clear plaques (50) was discarded. (The object of this was to keep the background of *c* I⁻ phage, which produce clear plaques, to a minimum.) Each plate was overlaid with 11.5 ml of Ø-dil, and plates were left on a rotating platform at 4°C with gentle agitation for 24 hours.

The liquid was decanted, pooled, and centrifuged at 6000X g for 10 minutes at 4°C to clear debris. 214.4 gm of CsCl (BRL, ultrapure) was dissolved in 268 ml of the phage supernatant (67 gm CsCl per 82 ml of Ø-dil yields a final volume of 100 ml with a density of 1.5 gm/ml), and this was loaded into 8 large heat-seal tubes (39 ml, Beckman) and centrifuged in a Beckman Ti70 rotor for 35 hours at 40,000 RPM, 4°C. The clear blue bands (*i.e.* phage) near the center of each tube were removed with a syringe, pooled, and an equal volume of Ø-dil was added. This was gently overlaid (~6 ml/ tube) a previously made CsCl step gradient consisting of 3 ml of CsCl in Ø-dil (density 1.4 gm/ml) which had been underlayed with 2.25 ml of CsCl in Ø-dil (density 1.6 gm/ml). These tubes were centrifuged in the SW 40.1 swinging bucket rotor for 2 hours at 26,000 RPM, 20°C. The phage bands were collected, pooled and 0.6 gm of CsCl was added per ml of liquid. This was placed in the bottom of a fresh tube (~6 ml/tube) and overlaid with 3 ml of 1.6 CsCl solution (84.2 gm CsCl/100 ml) followed by 3 ml of 1.4 CsCl solution (51 gm CsCl/100 ml). The tubes were centrifuged as above and highly purified phage particles were collected. Tris-HCl pH 8 was added to a final concentration of 50 mM, and EDTA to a final concentration of 20 mM. An equal volume of deionized formamide was added, and the mixture was allowed to stand at

room temperature for 3 hours (Davis *et al.*, 1980). One volume (*i.e.* equal to the volume of formamide) of water was added, and then 6 volumes of ethanol. The DNA was precipitated at -20°C for several hours, and then collected by spooling onto a glass rod. The precipitate was rinsed several times in 70% ethanol, dried, and resuspended in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Total yield of DNA was 3.4 mg, as determined by absorbance at 260 nm. 250 µg of this was cut with *Eco* RI (New England Biolabs), extracted with phenol, chloroform, and ether, precipitated with ethanol, and the entire procedure repeated to ensure that all DNA was cut to completion. The cut DNA (final yield 150µg) was resuspended in 200µl water (final concentration 750 µg/ml), heated to 70°C for 5 minutes, incubated at 42°C for 30 minutes (to anneal cohesive ends of the arms) and chilled on ice. ATP was added to 1mM and DTT to 10mM, and the arms were frozen in 25 µl aliquots at -70°C.

c) Extracts for In Vitro Packaging of Lambda Phage. The method of preparation of these two extracts is similar to that described by Maniatis *et al.* (1982). For the freeze-thaw lysate, an overnight culture of the bacterial strain BHB 2688 (obtained from M. Davis, Stanford) was grown in 25 ml of LB at 30°C. Three flasks, each containing 400 ml of LB were innoculated with 2 ml of the overnight culture, and these were grown at 30°C with vigorous aeration until the OD_{600} reached 0.350 (about 3.75 hours; the density was monitored closely). The cultures were quickly transferred to a shaking water bath at 45°C, and vigorous shaking maintained throughout this transfer and 15 minute heating period. The flasks were then returned to the air incubater (now warmed to 39°C), and allowed to incubate for a further 3.75 hours with vigorous shaking. Cells were transferred to pre-chilled centrifuge bottles, and pellets collected by centrifugation (Sorvall GSA rotor, 5000 RPM x 10 minutes, rotor and centrifuge pre-cooled to 4°C). The medium was decanted and the remaining liquid quickly removed with a pasteur pipet. Pellets were rapidly but gently resuspended in 1.5 ml

10% sucrose, 50 mMTris-HCl pH 8.0 (on ice), and the cell suspension transferred to a pre-chilled 50Ti ultracentrifuge tube; all bottles were rinsed with a second 1.5 ml aliquot of sucrose/Tris and this was added to the first. 150µl of lysozyme solution (2mg/ml lysozyme in 250 mM Tris-HCl pH 8.0) was added to the bacterial suspension, the tube was quickly flicked twice to mix the contents, and then snap frozen in liquid nitrogen for 5 min. The liquid was thawed on ice (this took about 1.5 hours), and centrifuged at 48,000 RPM for 1 hour at 4°C (50Ti rotor and ultracentrifuge were pre-chilled to 4°C by pre-running the empty rotor at low speed). The supernatant was stored on ice water and quickly aliquoted (30 µl aliquots) into pre-chilled Eppendorf tubes which were immediately snap frozen in liquid nitrogen. This process required two people: one to fill the tubes and the other to close them and place them in the nitrogen. The Eppendorf tubes were pre-chilled by placing them in an ice bucket which was one-third full of dry ice. The tubes were separated from the dry ice by two layers of Saran wrap to prevent over-chilling. (If the tubes are too cold, the Eppendorf pipet tip one is using to dispense the extract will freeze and become plugged.) Frozen extracts were transferred from the liquid nitrogen to -70°C for long term storage.

For the sonicated extracts the bacterial strain MMS A345 (obtained from D. Denny, Stanford) was used. A protocol for packaging using this single strain has been developed by F. Stahl (Eugene, OR). For our purposes the sonicated A345 extracts were combined with the freeze-thaw (packaging protein donor) extracts described above to obtain high packaging efficiencies. The innoculation and growth of the A345 cells was identical to that described above for BHB2688, except that the cells grew faster at 30°C so that with a 2 ml innoculum an OD_{600} of 0.40 was reached in 3 hours. The cells were then induced at 45°C as described, and incubated at 39°C for 2.25 hours. Cell pellets were collected and drained as for the freeze-thaw extract, and then resuspended in 6.5 ml of ice cold Sonic buffer (20mM Tris-HCl, 1mM EDTA, 5mM

2-mercaptoethanol). The bacterial suspension was transferred to a 30 ml cetrifuge tube (on ice water), and sonicated in short bursts (5 sec on, 25 sec off) using a Sorvall sonicator with the small tip at full power (A.R. Morgan). The suspension cleared (from bottom to top) after ~30 cycles of sonication, and was centrifuged in the Sorvall SS34 rotor at 10,000 RPM for 10 minutes (rotor and centrifuge pre-cooled to 4°C). 1.1 ml of ice cold Packaging buffer (6mM Tris-HCl pH 8.0, 50mM spermidine, 50mM putrescine, 20mM $MgCl_2$, 30mM ATP, and 30mM 2-mercaptoethanol) was added directly to the tube and mixed gently, taking care to avoid dislodging the pellet. The liquid was then quickly aliquoted into pre-chilled tubes (45µl/tube), snap frozen, and stored as described above. For packaging, one tube of each extract was thawed on ice, and as soon as it was liquid the freeze-thaw lysate was quickly added to the sonicated extract. The DNA was quickly added to the combined lysates and mixed gently and quickly with an Eppendorf pipet tip (with the tube still on ice), and the tube was left at room temperature for 1.5 hours. These extracts gave 2.5-2.9 x 10^8 pfu/µg of lambda phage DNA.

2. *Testing Enzymes and Buffers for cDNA Cloning*

The cloning of cDNA libraries involves a large number of complex enzymatic reactions, and each step in the process must be efficient if the overall process is to succeed. However, during the cloning process itself it is difficult or impossible to know if certain of the enzymatic steps have proceeded efficiently. In essence, one begins with RNA and finishes with a library of clones, and if the process does not succeed there is no simple way to discover which step has failed. One solution to this problem is to break the cloning process down into a number of discrete steps, and to devise an independent experiment to test the enzymes and buffers used in each step. For the cDNA cloning described in this thesis the steps are: 1) reverse transcription of RNA; 2) addition of homopolymer 'dG-tails' to the 3' end of first strand cDNA; 3)

oligo dC priming and second strand synthesis; 4) *Eco* RI methylation; 5) addition of *Eco* RI linkers, and cleavage with *Eco* RI; 6) size selection of cDNA on low melting point agarose gels; 7) ligation into vector; and 8) packaging of the phage recombinants. The experiments used to test the enzymes and buffers used in a number of these steps are described below, along with other relevant comments.

a) Reverse Transcription of RNA. This is one of the few steps whose progress can be assessed directly during the cloning process, since a low concentration of [alpha-^{32}P]dCTP is included in the reaction, and incorporation of the label into DNA can be monitored. Inclusion of a ribonuclease inhibitor improves yields significantly. The RNA should be intact as assessed by denaturing gels; the 28S band should be about twice as bright as the 18S band since they are present in equimolar amounts if the RNA is not degraded, and there should be a continuous smear of RNA throughout the gel to sizes considerably larger than 28S. The RNA should translate well *in vitro* if it is clean and intact.

b) Homopolymer Tailing Reaction. This reaction is extremely important if full length cDNA is to be generated. The addition of dG residues to the 3' ends of the first strand products eliminates the need to accurately control the lengths of the homopolymer tails added. This is because the secondary structure of the poly dG stretch causes the dG tailing reaction to terminate at useful 'tail' lengths (12-14 bp), and if the enzyme is sufficiently active > 95% of the ends will be extended. A simple, non-radioactive method to compare various batches of the enzyme terminal transferase and to test the tailing buffers involves cutting the *Pst* I insert out of a convenient plasmid and purifying the products by spermine precipitation. Aliquots of the resultant fragments are tailed with TTP or dCTP, and the various tailed preparations are compared with the untailed fragments on a 1% agarose gel. The tailed fragments will be shifted upward (*i.e.* have a lower mobility) in the gel compared to untailed fragments, and the amount

of shift will depend on the length of the tail and hence on the quality of the enzyme. The reason that dGTP is not used in this assay is that the short dG tails would not cause a visible mobility shift in the fragments. The addition of dC tails is generally more difficult (*i.e.* the resultant tails are not as long) than the addition of T tails, and therefore the dC reaction allows greater discrimination between various preparations of enzyme.

c) Second Strand Synthesis. Preparations of the enzyme DNA polymerase I used in this reaction are usually of high quality and generally do not need to be tested. The enzyme should rapidly and efficiently incorporate ^{32}P-dNTPs into ds DNA during nick translation. It should actively promote filling in of 5' protruding restriction ends, and efficiently remove 3'-protruding restriction ends.

d) Methylation of ds-cDNA. *Eco* RI methylation usually proceeds with high efficiency to completion. To test this reaction, a few µg of a DNA fragment which contains a number of *Eco* RI sites (*e.g.* lambda DNA) is divided into three aliquots. One aliquot is methylated and then cut with *Eco* RI, the second is simply cut with *Eco* RI, and the third is left as a control. Methylation should completely block cleavage at any of the *Eco* RI sites which would otherwise be cut to completion, and this will be apparent when the three aliquots of DNA are run on a gel and compared.

e) Addition of *Eco* RI linkers and cutting with *Eco* RI. A simple, non-radioactive test of this reaction is the following. A few µg of pBR322 is cut with *Hae* III to generate a series of blunt-ended fragments. The cut DNA is purified by spermine precipitation, and one-third of the DNA is removed and saved. The remaining DNA is added to phosphorylated *Eco* RI linkers and the ligation reaction allowed to proceed. The reaction is heated to 65°C for 10 minutes to inactivate the ligase, and then one-half of the remaining DNA is cut to completion with *Eco* RI. The three aliquots of DNA (*Hae* III cut, *Hae* III cut- *Eco* RI linker ligated, and *Hae* III cut- *Eco* RI linker ligated, *Eco* RI cut) are then run on a 2% agarose gel and compared. Ligation of linkers should

make the *Hae* III bands disappear into a continuous smear of DNA, and cutting with *Eco* RI should make the pattern of *Hae* III bands reappear, but with the bands all shifted to a slightly larger size since they now have *Eco* RI linkers on both ends of each fragment (see Figure 2).

f) Size Selection of ds-cDNA on Low-Melting Point Agarose Gels. It is extremely important to purify cDNA of the desired size class away from shorter, incomplete cDNA molecules and linkers. If this step is not done, a large amount of time is subsequently wasted 'chasing down' many clones which contain only a tiny fraction of the cDNA molecule of interest, since these shorter inserts are more likely to be ligated into the vector if they are present during the ligation. Several methods have been used with success to 'size select' cDNA, including column chromatography (Huynh *et al* ., 1985) as well as direct preparative electrophoresis. In the cloning described in this thesis, the *Eco* RI cut cDNA was run on a 1.4% low melting point agarose gel (BRL), and cDNA of the desired size (> ~800 bp.) was cut directly out of the gel and extracted using the 'CETAB' method (Langridge *et al*., 1980). The ethidium bromide stained gel gives one the opportunity to visualize directly the size and quantity of cDNA produced, and recovery from the gel can be followed by ^{32}P counts, since the first strand of the cDNA is labelled.

g) Ligation of cDNA into Lambda Vector Arms. Once the ds-cDNA is completed, size selected, and purified, it is prudent to do a number of pilot reactions whereby a small portion (~ 5%) of the cDNA is ligated into vector arms, and the resultant phage is packaged and titred. This ensures that the final steps in the cloning process are working, and means that if they fail, most of the cDNA is still available for subsequent cloning without returning to the very beginning of the process. To measure the very small volumes required, a 10 µl micro syringe (Glenco) with 0.1 µl gradations was used, and 'disposable tips' were made by cutting ~6 cm. lengths of teflon tubing

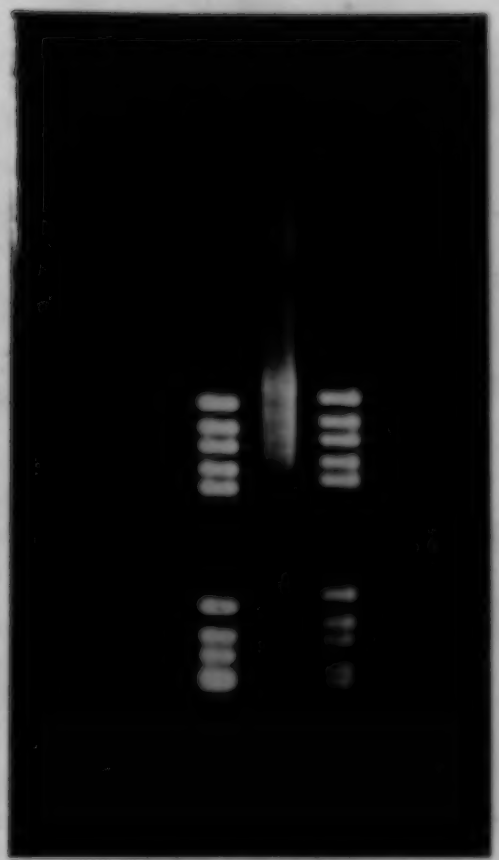


FIGURE 2. Test of *Eco* RI linker ligation. Ten µg of pBR322 plasmid DNA was cut to completion with the restriction enzyme *Hae* III, spermine precipitated, and resuspended in 10 µl of 50 mM Tris-HCl (pH 7.5), 10 mM magnesium chloride. Two µl of *Eco* RI linkers (1.0 OD/100µl) were kinased in a 10 µl reaction volume, and 6 µl of the cut plasmid DNA was added to the kinased linkers. 0.15 µl of ATP (100 mM) and 0.4 µl of T4 DNA ligase (400 U/µl, New England Biolabs) was added and the reaction was incubated at 14°C overnight. The ligase was inactivated by heating to 65°C for 5 minutes, and 0.2 µl of 5 M NaCl was added to a 8.5 µl aliquot of the ligated mixture. This was cut with *Eco* RI and heated to 65°C for 10 minutes. The three samples of DNA were run on a 2% agarose gel. The lanes are (from left to right): 1- pBR322 cut with *Hae* III; 2- *Hae* III cut DNA with *Eco* RI linkers liagated; 3- *Hae* III cut DNA with *Eco* RI linkers ligated, now recut with *Eco* RI.

(Clay Adams # 7406) and fitting these over the blunt end of a shortened needle. The phage lambda gt10 has a single *Eco* RI site within the the *c* I repressor gene; phages in which this gene is interrupted will produce clear plaques (*c* I⁻), whereas phages in which the *c* I gene is intact will produce cloudy plaques. During the pilot ligations the amount of cDNA ligated per µg of *Eco* RI cut lambda gt10 arms was systematically varied, so that the number of 'clear' plaques which resulted because cDNA inserts interrupt the *c* I gene could be optimized. If *Eco* RI cut arms were simply ligated and packaged, a certain frequency of clear plaques resulted. This background of *c* I⁻, clear plaques (0.1-0.2% in these experiments) appeared to be unavoidable, and the challenge during the pilot ligations was to use a minimum amount of cDNA to increase the frequency of *c* I⁻ (*i.e.* recombinant) phage significantly above the background. The optimum ratio was determined to be 12 ng ds-cDNA per 1 µg of phage arms, and these ligations generally resulted in a frequency of clear plaques of 0.4-0.5%. This meant that when these libraries were passaged through an *E. coli hflA* mutant strain (which allow only *c* I⁻ phage to form plaques, Huynh *et al* ., 1985; Hoyt *et al*., 1982), only one-half to two-thirds of the resultant plaques actually contained cDNA inserts. In theory the frequency of clear plaques should be determined by plating the library on a strain such as KM392, and counting the number of clear and cloudy plaques. However, for practical reasons the number of cloudy plaques was determined by plating the library on KM392, and the number of 'clear' by plating on a C600 *hflA* strain (obtained from T. Huynh, Stanford).

Once an enzyme or buffer successfully 'passed' one of the various tests outlined above it was kept for personal use only, and used exclusively for constructing cDNA libraries. (*i.e.* The John Elliott 'School of Paranoia.')

3. *Construction of cDNA Libraries*

Complete details of the steps taken during the cloning of a cDNA library are given in this section. Note that all Eppendorf tubes have been siliconized, rinsed, and autoclaved prior to use.

a) First Strand cDNA Synthesis. 17.5 µl of water, 5 µl of 10x RT salts, 10 µl of all four dNTPs at 5 mM each (Pharmacia), 5 µl of oligo dT_{12-18} (1 mg/ml, Collaborative Research), 0.5 µl of Vnx (0.2 M, made as described above), 5 µl of poly A^+ RNA (2 mg/ml), and 2 µl [alpha-^{32}P]dCTP (used on the day of arrival, 3200 Ci/mmol, New England Nuclear) were combined in that order in a siliconized tube on ice. 3.5µl of AMV Reverse Transcriptase (17U/ml, Life Sciences) was added, the contents were mixed by vortexing and incubated at 41°C for 1 hour. A second aliquot of 2.0 µl Reverse Transcriptase was added and the reaction incubated for a further 90 minutes at 41°C. The 10x RT salts consisted of 1.0 M Tris-HCl (pH carefully adjusted to 8.3 at 41°C), 0.5 M KCl, 0.1 M $MgCl_2$, and 0.1 M DTT. The reaction was terminated by the addition of 2 µl of 0.5 M EDTA. 10 µl of 80% glycerol, 0.2% bromophenol blue was added, and the mixture was loaded onto the submerged bed of a 0.3cm x 23cm Biogel A-5m (Biorad) column equilibriated in 1 mM Tris-HCl pH 8.0, 0.01 mM EDTA. One drop fractions were collected from the column, counted, and fractions containing the initial peak of radioactivity (excluding the two fractions just before the 'trough') were pooled in a siliconized tube and lyophilized overnight.

b) Tailing Reaction. The dried cDNA was resuspended in 72 µl of water, counted, and yield of first strand material calculated. This varied between 4 and 6 µg. 8 µl of sodium cacodylate (1.0 M, pH 7.0), 4 µl dGTP (20 mM), 1.6 µl $CoCl_2$ (100 mM), and 3 µl of terminal transferase (21U/µl, Ratliff Biochemicals, Los Alamos) was added to the cDNA, and the reaction was incubated at 37°C for 30 minutes. A second aliquot of 3 µl of enzyme was added, and incubation continued for an additional 30 minutes, followed by a third addition of enzyme (2 µl) and a further 30 minute incubation (total

incubation was 90 minutes at 37°C). 4 µl of EDTA (100 mM), and 92 µl of TE (10 mM Tris-HCl pH 7.4, 1 mM EDTA) were added to the reaction and it was placed in a 70°C water bath for 5 minutes to inactivate the terminal transferase.

c) Second Strand Synthesis. 0.8 µl of RNAse A (1 mg/ml, boiled 2 x 15 minutes, Boehringer-Mannheim), and 2.4 µl of oligo dC_{12-18} (1 mg/ml, Collaborative Research) were added to the tailed cDNA, and the tube heated in a boiling water bath for 1 minute, then chilled on ice. 13.2 µl of $MgCl_2$ (100 mM), 4.8 µl of all four dNTPs (at 5 mM each), and 4.0 µl of DNA polymerase I (200 U/µl, New England Biolabs) were added to the tube on ice, and the reaction was incubated for 21 hours at 14°C. The polymerase was inactivated by the addition of 9.6 µl of EDTA (500 mM), and heating in a 70°C waterbath for 10 minutes.

d) Methylation, Organic Extraction and Spermine Precipitation. The tube was chilled on ice, and 5 µl of S-adenosyl-L-methionine (SAM, 1 mM, Pharmacia), and 2 µl of *Eco* RI methylase (40U/ml, New England Biolabs) were added. The methylation reaction was incubated at 37°C for 30 minutes. 200 µl of phenol-chloroform (1:1, equilibriated with TE) was added to the tube, it was vortexed and centrifuged and the aqueous phase placed in a second siliconized Eppendorf tube. The remaining organic phase was back extracted twice with 100 µl of TE each time and the TE phases pooled with the original aqueous. The recovery of DNA into the aqueous phase was monitored by ^{32}P counts, and was usually ≈90%. The pooled aqueous phases were extracted 3 times with 800 µl of water saturated ether (diethyl, Fisher), and the last traces of ether were removed by heating to 65°C for 5 minutes and then blowing nitrogen gas over the surface of the liquid for 30 minutes (R. McElhaney's lab). The original aqueous volume (determined by marking the tube prior to heating) was restored by addition of water as necessary. The DNA was precipitated by the addition of 15µl of spermine (100 mM), vortexing, and storing the tube on ice for 30 minutes. The DNA was

pelleted by spinning 10 minutes in the microfuge (Brinkmann), and recovery (usually >85%) was monitored by ^{32}P counts. 1 ml of 75% ethanol, 0.3 M sodium acetate, 10 mM magnesium acetate was added to the DNA pellet, and this was left on ice for 1 hour with occasional vortexing (this removes the spermine). The DNA was centrifuged, and the pellet was washed with 1 ml of 70% ethanol and stored for 15 minutes on ice. The pellet was again collected by a 10 minute centrifugation, and finally dried for 5 minutes in the speed-vac.

e) Phosphorylating and Ligating *Eco* RI Linkers. 6 µl of *Eco* RI linkers (1.0 A_{260} unit/100 µl water, New England Biolabs #1004), 6 µl of 0.5 M Tris-HCl pH 7.5/ 0.1 M $MgCl_2$, 1µl of ATP (100 mM), 3 µl of DTT (200 mM), and 44 µl of water were combined in a siliconized Eppendorf tube. 0.6 µl of T4 polynucleotide kinase (6 U/µl, Pharmacia) was added, and the reaction was incubated at 37°C for 20 minutes. An additional 0.6 µl aliquot of enzyme was added, and the reaction incubated for a further 20 minutes, followed by a third aliquot and a third 20 minute incubation. The phosphorylated linkers were added directly to the dried ds-cDNA pellet (from d] above) on ice, and 1.2 µl of ATP and 1.4 µl of T4 DNA ligase (400 U/µl, New England Biolabs) were added to the tube and the contents mixed thoroughly to ensure that the DNA pellet was dissolved. The ligation reaction was incubated at 14°C for 24 hours, and then heated to 65°C for 10 minutes to inactivate the ligase.

f) *Eco* RI Cleavage and Size Selection of cDNA. The DNA was cooled on ice, and 1.2 µl of NaCl (5.0 M) and 2.3 µl of *Eco* RI (20U/µl, New England Biolabs) were added and the reaction was incubated at 37°C for a total of 90 minutes. Two further aliquots of 2.3 µl *Eco* RI were added at 30 minute intervals during the reaction period. The *Eco* RI was inactivated by heating the reaction to 65°C for 10 minutes, 10 µl of 10X loading dye (Maniatis *et al.*, 1982) was added, and the cDNA was run on a 1.4% low melting point agarose gel (BRL) and size selected as described in the previous

section. DNA of the desired size was extracted from the gel using the CETAB procedure (see above), and the DNA present in the final aqueous phase (0.2 M NaCl) was precipitated with 3 volumes of ethanol (-70°C). The DNA was collected by centrifugation (15 minutes in a microfuge at 4°C [*i.e.* in the cold room]) and the pellet was dried and resuspended in 200 µl of KCl (100 mM). The DNA was precipitated a second time by the addition of 6.0 µl of spermine (100 mM), and the precipitation and washing of the pellet was done as described above. The final dried pellet was resuspended in 20 µl of water, counted to determine the yield, and stored at -70°C. The yield was usually about 400 ng of ds-cDNA (greater than ≈800 bp in length) for the reaction conditions described above.

g) Ligation of cDNA into Phage Vector Arms and Plating of Libraries. A typical pilot ligation reaction contained 2 µl of *Eco* RI cut lambda gt10 arms (1.5 µg, prepared as described above), 1 µl of cDNA (12-25 ng), 0.1 µl of 0.1 M $MgCl_2$, and 0.3 µl of T4 DNA ligase (400U/ml, New England Biolabs). The reaction was incubated for 20-24 hours at 14°C, then chilled on ice and packaged by the addition of both packaging extracts to the tube containing the ligated DNA. A control ligation containing water instead of cDNA was always done in parallel. If the pilot ligation gave a satisfactory frequency of 'clear' plaques (as discussed above), it was scaled up directly and the majority of the cDNA ligated and packaged, using a maximum of 1.5 µg of input phage arms per packaging reaction. Libraries typically contained 5-10 x 10^6 *c* I⁻ phage, about half of which contained cDNA inserts. 2 x 10^6 *c* I⁻ phage were plated out on the C600 *hflA* strain using 20 large (150 mm) plates. These were grown at 37°C until the plaques were confluent (6-7 hours), then the plates were overlaid with 12 ml of Ø-dil and the phage allowed to elute at 4°C on a rotating platform with gentle agitation. The liquid from the plates was pooled in a sterile glass bottle, 5 ml of chloroform was added, and the phage stock was stored at 4°C for several years. This represents the

amplified library. At the time of packaging, the original library was also plated on the C600 *hflA* strain using about 10 large plates (50,000 *c* I⁻ phage per 150 mm plate), and these plates were screened directly for cDNAs of interest.

E. ISOLATION AND CHARACTERIZATION OF CLONES FROM cDNA LIBRARIES

1. *Screening Libraries and Plaque Purification of Positive Clones.*

The cDNA libraries in phage vectors were screened using ^{32}P labelled DNA probes. The original libraries were plated as described above, and incubated at 37°C until the plaques were ≈0.5 mm in diameter. The plates were then chilled in an upright position for several hours at 4°C, and dry nitrocellulose circles were placed on each plate and allowed to remain until they were completely translucent (*i.e.* 'wet through'; this required about 1 minute). The position of the filters on the plates was recorded by poking through the filter in 6 or 7 places with the very sharp end of a drawn out pasteur pipet which had been dipped in india ink (Sanford, not Speedball). The filter was gently 'peeled' off the surface of the plate using millipore forceps, soaked in denaturing solution (1.5 M NaCl, 0.5 M NaOH) for ≈45 seconds, and transferred to neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl pH 8.0) for 2-3 minutes. A second 'duplicate' filter was placed on each plate and left for several minutes, marked with india ink in the same spots as the first filter, and denatured and neutralized as above. All filters were air dried, then baked 2 hours at 80°C in a vacuum oven. Filters were prehybridized, hybridized to labelled oligonucleotide probes, and washed using the conditions described in a previous section.

Washed filters were air dried and scotch taped to larger sheets of 3mm paper. The 3mm paper was in turn labelled with a number of characteristic marks using ^{32}P labelled india ink, and the filters and paper were wrapped in saran wrap and exposed to

X-ray film. The developed film could be aligned with the original paper because of the labelled india ink marks, and the position of the 'alignment' holes in the nitrocellulose filters could then be determined and marked on the film. In this way positive clones could be located on the original plate, and signals from duplicate filters could be aligned and compared. In general, clones were only considered true positives if they gave a signal on both of the duplicate filters, using a different oligonucleotide to probe each of the filters. The general region surrounding a 'true positive' signal was picked from the original plate using the fat end of a sterile pasteur pipet, and the agar plug was placed in 1 ml of Ø-dil, vortexed, and left for several hours. The resulting phage stock was titred on smaller (90 mm) petri dishes, and a single nitrocellulose filter was replicated from a plate containing 500-1000 plaques. This filter was hybridized to one of the labelled probes, and a positive plaque identified and 'pulled' from the smaller plate using the sharp end of a pasteur pipet. This plug was again placed in 1 ml of Ø-dil, vortexed, and phage particles allowed to elute. This final phage stock was again titred, and a nitrocellulose filter was replicated from a plate containing 10-30 well isolated plaques. This filter was hybridized to the second labelled probe, and one or two well isolated, positive plaques were finally obtained. These were placed in 1 ml of Ø-dil, 1 drop of chloroform was added, and the phage stock was stored at 4°C.

2. *Preparing cDNA Inserts and Subcloning into Plasmid Vectors.*

Small scale liquid cultures of (plaque purified) phage clones were grown, and phage DNA isolated essentially as described in Maniatis *et al.* (1982). 0.5 ml of a fresh plating stock of *E. coli* C600*hflA* (grown in LB with 0.2% maltose to an OD_{600} of ≈1.5, pelleted, and resuspended in one-half volume of Ø-dil) was combined with 0.4 ml of phage stock (containing 1 plaque in 1 ml of Ø-dil) in a 50 ml disposable polypropylene tube, and the mixture was incubated at room temperature for 20 minutes. During this period a small hole was melted in the side of the tube, very near the top,

using a paper clip heated in a bunsen burner. This hole allows the culture to be aerated even when the top of the tube is screwed on tightly. Following the 20 minute incubation, 7 ml of LB (supplemented with 10 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$) was added to the tube, and the tube was shaken very vigorously (325 RPM) at 37°C in an air shaker. The cultures generally became quite cloudy by 3-4 hours (as the bacteria grew to a high concentration), and then fairly abruptly 'cleared' in the next 30-60 minutes, leaving only fine strands of bacterial debris (as the bacteria lysed and released the phage). If the culture cleared too quickly, an additional 1ml of plating stock was added to the tube and the incubation continued. If the culture did not clear, it was left for 6-7 hours (or occasionally overnight). Cultures which did not eventually clear were discarded, and the culture process was repeated using a higher ratio of phage to bacteria in the original innoculum. A drop of chloroform was added to cultures which had cleared, and they were incubated a further 15 minutes with shaking. The cultures were then centrifuged to remove bacterial debris, incubated with RNase A and DNase I, and phage particles were precipitated with polyethylene glycol and DNA extracted as described in Maniatis *et al.* (1982). The phage DNA was cut to completion with *Eco* RI, treated with RNase A, and run on a 0.8% low melting point agarose gel (or alternately on a 3.5% acrylamide gel). The short *Eco* RI band representing the cDNA insert was cut out of the gel, and DNA extracted using standard methods (phenol extraction in the case of low melt agarose, 'crush and soak' in the case of acrylamide; both are described in Maniatis *et al.* [1982]). The purified cDNA insert was spermine precipitated, and ligated into *Eco* RI cut, phosphorylated pUC 13. The recombinant plasmid was used to transform *E. coli* HB 101, using frozen competent cells prepared as described by Hanahan (1985). The cDNA inserts in the plasmid were then mapped with a number of restriction enzymes, and larger quantities of the purified insert were obtained for use as probes or for futher subcloning into M13 phages or expression

vectors.

F. ISOLATION OF PMA-INDUCED cDNA CLONES FROM EL4

Isolation of PMA induced cDNA clones was accomplished in two steps. A subtracted library was first constructed, and this was in turn screened with subtracted cDNA probe. Methods for generation of both the library and the probe were adapted from unpublished protocols developed by M. Davis (Stanford). In each case RNA from PMA-treated EL4 cells was used to generate cDNA, and this was hybridized to mRNA from untreated EL4 cells in order to enrich for those sequences present only in the 'induced' (PMA treated) cells.

1. *Construction of Subtracted cDNA Library.*

First strand synthesis was accomplished in a mixture containing 250 µl water, 100 µl actinomycin D (1 mg/ml in water, Sigma), 100 µl of 10x RT salts (described above), 200 µl of all four dNTPs (at 5 mM each), 100 µl oligo dT_{12-18} (1 mg/ml), 10 µl of Vnx (200 mM), 100 µl poly A^+ RNA (2 mg/ml, from PMA treated EL4 cells), and 40 µl [alpha-^{32}P]-dCTP (used on the day of arrival, 3200 Ci/mmol, New England Nuclear). 60 µl AMV reverse transcriptase (17U/µl, Life Sciences) was added, and the reaction incubated at 41°C for 1 hour. A second aliquot of 40 µl of reverse transcriptase was added, and after a second hour of incubation a third aliquot of 20 µl was added, and the reaction incubated for a further hour. The mixture was chilled on ice, and 10 µl of 10 M NaOH was added to the tube and the contents quickly mixed. The tube was heated to 70°C for 20 minutes to hydrolyze the RNA, then chilled on ice, and the contents neutralized by adding 100 µl of 1M Tris-HCl pH 7.5, and then ≈100 µl of 1 M HCl (the pH was monitored by placing 3 µl aliquots on tiny strips of pH paper). The first strand products were heated to room temperature, 12 µl of 10% SDS was added, and they were loaded on 12 identical ≈1.8 ml G-50F columns equilibriated

in 1mM Tris-HCl pH 8.0, 0.01mM EDTA (100 µl per column). The columns were 'run' in parallel, and 100 µl fractions were collected from each. (This large number of columns were run simply as a matter of convenience, since the smaller columns have a high flow rate, and columns with this configuration had already been calibrated by the author.) The excluded peak of radioactivity (usually fractions 7-13) was pooled from each of the columns, and these were in turn pooled and concentrated with sec-butanol to a volume of 700µl. This was frozen on dry ice and lyohillized overnight in a siliconized Eppendorf tube. At the same time, 200 µg of poly A^+ RNA (2 mg/ml, from 'untreated' or control EL4 cells) was frozen and lyophilized.

The RNA was dissolved in 40µl of water, and this was transferred to the tube containing the dried cDNA. The cDNA was dissolved by vortexing and heating to 42°C. The tube which originally contained the RNA was rinsed with 10µl of 2 M phosphate buffer (PB, pH=6.8, containing equimolar amounts of Na_2HPO_4 and NaH_2PO_4), and with a further 5µl of PB, and these were transferred to the tube containing the DNA and RNA. 1µl of 10% SDS and 0.5µl of EDTA (0.1 M) were added to the mixture, and it was stirred gently with a pipet tip (vortexing or pipetting following the addition of SDS causes an impossible number of bubbles). The mixture was taken up in a 50µl siliconized glass capillary tube using an accupet adapter fitted onto the end of an ordinary P200 Pipetman, and both ends of the tube were heat sealed. The tube was heated for 30 seconds in a boiling water bath, and then incubated at 68°C for 6 hours. The total R_0t for this hybridization was ≈1600 moles of nucleotides/litre x seconds, which was calculated using the formula R_0t=(µg/ml RNA)x(0.07)x(hours of hybridization), based on a phosphate concentration of 0.5 M and a temperature of 68°C, and using a correction factor of 5.82 (Britten *et al.*, 1974). Following hybridization the ends of the tube were broken using a triangular file, and the contents of the tube were diluted into 1 ml of 0.12 M PB, 0.1% SDS (HAP buffer). This was applied to 1.5 ml

of hydroxyapatite (Biorad, DNA grade) which had been previously equilibriated with HAP buffer in a water jacketed column at 60°C. The liquid was gently forced through the hydroxyapatite by applying a low intermittent pressure (~ 5 PSI) to the column, using a one-hole stopper which fit into the mouth of the column and which was connected via a flexible hose to the general air pressure line. (The valve in the air pressure line was just open, so that a very gentle stream of air passed out of the stopper continuously.) The column was washed with 7 additional 1ml aliquots of HAP buffer (at 60°C), and with each aliquot pressure was applied to the column until the hydroxyapatite was almost dry. The eight resulting 1 ml fractions from the column were collected in separate tubes, and aliquots of each tube were counted. The majority of the single stranded material (*i.e.* ^{32}P-counts) was in the first three fractions, and these were pooled, concentrated with sec-butanol to a volume of 0.7 ml, frozen in dry ice and lyophilized overnight. The double-stranded material was eluted from the hydroxyapatite by heating the column to 97°C and washing it with ten 1 ml aliquots of HAP buffer at 97°C. This material was pooled and counted, so that the total amount and percent of counts bound could be calculated.

The dried single-stranded cDNA was dissolved in 200 µl of water and passed over a ~1.8ml (pasteur pipet) G50-F column equilibriated in 1mM Tris-HCl pH 7.5, 0.01mM EDTA. 100 µl fractions were collected, and fractions 7-15 inclusive were pooled, frozen, and lyophilized. (The 'excluded' peak of radioactivity did not appear to be well defined [there was essentially no trough], and these fractions were chosen because they were likely to include all of the cDNA of a reasonable size.) The dried cDNA was redissolved in 52 µl of water and 10 µl of 80% glycerol; 0.2% bromophenol blue was added, and the material was loaded onto the submerged bed of a Biogel A-5m column (0.3 cm x 23 cm, equilibriated in 1mM Tris-HCl pH 8, 0.01 mM EDTA). All subsequent steps in the cDNA cloning process, including the 3' dG tailing

are identical to those described in the previous section.

2. *Preparation of Subtracted cDNA Probe.*

First strand synthesis was accomplished in a siliconized Eppendorf tube containing 100 µl [alpha-^{32}P]-dCTP (used on the day of arrival, >3000 Ci/mmol, New England Nuclear), 25 µl of 10x RT buffer, 10 µl dATP(20 mM), 10 µl dGTP (20 mM), 10 µl TTP(20 mM), 5µl dCTP (2mM), 20 µl oligo dT_{12-18} (1 mg/ml), 2.5 µl Vnx (200 mM), 25 µl actinomycin D (1 mg/ml), and 20 µl of RNA (2mg/ml, poly A^+ from PMA-treated cells). A total of 30 µl of AMV reverse transcriptase (22 U/µl, Life Sciences) was added to the reaction (in aliquots of 15, 10, and 5 µl at hourly intervals), which was incubated at 41°C for a total of 3 hours. The RNA was hydrolyzed by adding 2.6 µl of 10 M NaOH and heating to 70°C for 25 minutes. The mixture was neutralized with 30µl of 2.0 M Tris-HCl pH 7.5 and ≈20µl of 1 M HCl (determined empirically by neutralizing a 'mock' tube containing the same RT buffer, base, and Tris-HCl). 3 µl of 10% SDS was added, the DNA was loaded on a G-50F column (equilibriated with 1mM Tris-HCl pH 8, 0.01 mM EDTA), and the excluded peak of radioactivity pooled, concentrated with sec-butanol, and lyophilized. Yields were typically about 12 µg, or 900 x 10^6 cpm of single stranded cDNA. 120 µg of poly A^+ RNA (from untreated EL4 cells) was lyophilized and then dissolved in 21.5 µl of water (containing 2 mM Vnx). The RNA solution was added to and used to dissolve the dried cDNA. The tube which had contained the RNA was rinsed with 7.5 µl of 2 M PB, and this was added to the mixture. 0.8 µl of 10% SDS, and 0.4 µl of EDTA (100 mM) were added, and the contents of the tube were taken up in a 50 µl siliconized capillary tube. The ends of the tube were heat sealed and it was heated in a boiling water bath for 30 seconds, and then incubated at 68°C for 6 hours. The total $R_o t$ for this hybridization was 1680 moles of nucleotides/litre x seconds, which was calculated using the formula $R_o t$=(µg/ml RNA)x(0.07)x(hours of hybridization), based on a

phosphate concentration of 0.5 M and a temperature of 68°C, and using a correction factor of 5.82 (Britten *et al.*, 1974).

The single stranded cDNA was separated from double stranded material on a hydroxyapatite column as described, and the fractions containing ss-cDNA were sec-butanol concentrated to a volume of 200 µl and run over a G50-F desalting column as in the previous section. Fractions 7-13 were pooled (100 µl fractions; again there did not appear to be a well defined peak of excluded radioactive material), and lyophilized. The entire hybridization process was repeated, except that 60 µg of lyophilized RNA was dissolved in 12 µl of water (with 2 mM Vnx), and 4 µl of 2 M PB, 0.4 µl of 10% SDS, and 0.2 µl of EDTA were added before the mixture was placed in a 20µl glass capillary. The ss-cDNA remaining after hydroxyapatite chromatography was concentrated, desalted, and fractions 7-13 pooled, counted, and used directly to probe nitrocellulose filters replicated from the subtracted library. Yields of subtracted probe were typically 25-40 x 10^6 cpm.

A variation on the basic method of making subtracted probe was used in later experiments. In this protocol, first strand cDNA was made using 10% as much [alpha-^{32}P]-dCTP as above (*i.e.* 10 µl instead of 100 µl, the remaining volume is replaced with water). The remainder of the subtraction process (through two cycles) was identical. The final ss-cDNA which was run over a desalting column was sec-butanol concentrated and then precipitated with ethanol and sodium acetate (RNA present in the DNA acts as carrier). The pellet was dissolved in 50 µl TE and it was stored at -70°C. When subtracted probe was required, a portion of the ss-cDNA was labelled using the random primer method as outlined in section C.2. For example, 25 µl of the DNA solution was added to 30 µl of water, and the mixture was placed in a boiling water bath for 4 minutes. The tube was cooled to 37°C, and 2µl of RNase A (2mg/ml), 20 µl of OLB, 4µl of BSA (2 mg/ml), 10 µl of [alpha-^{32}P]-dCTP (>3000

Ci/mmol, New England Nuclear), and 4 µl (20 units) of 'klenow' added. The reaction was incubated for 6 hours, boiled for 4 minutes, cooled to 37 °C, and an additional 50 µl of water, 20 µl of OLB, 8 µl of BSA, 20 µl of [alpha-^{32}P]-dCTP, and 4 µl of klenow added and the reaction incubated overnight. This method generated roughly equivalent amounts of probe as did the previous one (in terms of total number of counts), but used considerably less radioactive nucleotide and left half of the material for labelling at a later time.

3. *Screening the Subtracted Library*

The subtracted library was plated out on 6 large (150 mm) plates, with ~50,000 plaques per plate, using the host C600 *hflA* . Replicate nitrocellulose filters were pulled from each plate, and one set was prehybridized in a 50% formamide solution as described. These were then washed 6 times in 4X SSPE, and placed in pairs (back to back) in 'baggies' which were then heat sealed, taking great care to exclude as much of the air as possible from the bag. The subtracted probe was added to a hybridization mix containing 1.0 M NaCl, 0.2 M PB, 2X Denhardts solution, 0.1% SDS, and 10% dextran sulfate, so that the final concentration of probe was 2-3 x 10^6 cpm/ml. About 3 ml of this probe solution was introduced into each of the pre-sealed bags, using a 22 gauge needle and 5 ml syringe. The remaining air in the bag was carefully excluded, and the 'puncture' hole sealed. These filters were hybridized at 65°C for 48 hours, and then washed as described in section C.3. Apparent positive clones were picked using the fine end of a sterile pasteur pipet, and placed in 1 ml of Ø-dil. Because the nitrocellulose shrinks during the baking and hybridization process, it was easiest to align positive signals with regions of the plate very near the 'alignment' marks on the filters, and it was primarily from these areas that plaques were picked. The stocks containing positive clones were individually replated at an appropriate density, and

rescreened once or twice as necessary (using a second preparation of subtracted probe) until positive clones were plaque purified.

G. CHARACTERIZATION OF PMA-INDUCED cDNA CLONES

1. *Grouping Clones Into Families, Subcloning, and Northern Analysis*

One hundred and thirty-six positive clones were initially picked and plaque purified. Phage stocks representative of the individual clones were placed in an ordered array in 96-well microtitre plates, and each array was replica plated multiple times onto fresh lawns of C600 *hflA* (0.2% maltose in both the top agar and the plates) using a 48-pin microtitre transfer device (Tyler Research Corporation). The resulting plaques were transferred onto nitrocellulose, to give a large number of filters each of which contained an ordered array of all of the sequences contained in the original 136 clones. Twenty clones of the 136 were chosen at random and in each case *Eco* RI inserts were purified and subcloned into pUC 13. cDNA inserts cleaved from the pUC subclones were obtained in slices of low melting point agarose and labelled directly using the random primer method. The labelled inserts were used to probe northern blots containing poly A^+ RNA from both PMA treated and untreated EL4 cells, to confirm that the clones represented PMA induced transcripts. The labelled inserts were also used simultaneously to probe a set of the filters containing DNA from the 136 original clones, so that the clones could be placed into a series of cross-hybridizing groups or families. This also served to confirm that the plasmid subclone was representative of the original phage clone. By subcloning those phage clones which did not hybridize initially, all of the original clones were rapidly placed in groups and representative members of each were available in plasmids for further study. A representative member (usually the clone with the longest cDNA insert) of each 'family' of induced clones was used to probe northern blots of total cellular RNA to re-confirm that the clone

represented a PMA induced transcript and to determine the effects of cyclosporin A and cyclohexamide on the expression of the transcript in EL4.

2. *Primary Sequence Analysis of PMA-Induced cDNA Clones*

DNA sequencing was accomplished using single stranded templates and the dideoxynucleotide chain termination method developed by Sanger *et al.* (1977). In general, the procedures used followed closely those described in the M13 Cloning and Sequencing Handbook (Amersham). The details of several procedures which have been changed or adapted are given below.

a) Subcloning into M13 and Generation of Deletion Mutants. Purified cDNA inserts were ligated into M13mp18 RF (Yanisch-Perron *et al.*, 1985) which had been previously cut with *Eco* RI and dephosphorylated using calf intestinal phosphatase (Boehringer Mannheim) as outlined in Maniatis *et al.* (1982). Transformations and plating were done in *E. coli* strains JM 101 or JM109 (the latter has the advantage of being *recA1* minus). Clones with inserts in either orientation were identified using the C-Test described by Messing (1983). For cDNA clones that contained an internal restriction site that was also present in the M13 mp18 polylinker, a direct method of generating deletion mutants was used. A small amount of the double-stranded RF form of the recombinant was obtained using the 'miniprep' procedure outlined in Maniatis *et al.* (1982). This DNA was cut with the relevant restriction enzyme, the restriction enzyme was inactivated, the digest diluted 10-20 fold, and ligated directly. Virtually all of the recombinant phage made in this way had deleted the portion of the cDNA insert between the restriction site in the polylinker and that in the insert. Other deletants were constructed by cutting the cDNA insert with relevant restriction enzymes, isolating the various fragments from acrylamide gels, and ligating into appropriately cut M13 mp18 or M13 mp19 vectors.

A generalized method for generating deletion mutants in M13 was adapted from

the protocol of Guo *et al.* (1983). A 'cassette' vector was prepared by cutting M13 mp18 RF DNA with *Bam* HI, filling in with dNTPs using the klenow fragment (of DNA polymerase I), and then cutting with *Eco* RI. The result was a vector with one blunt end, and one *Eco* RI 'sticky' end; the vector DNA was also dephosphorylated so that it was unlikely to simply close on itself. 25 µg of plasmid (pUC or pGEM1, Promega Biotec) containing the cDNA insert to be deleted was cut with *Hind* III (or *Bam* HI), phenol-chloroform extracted, and precipitated with sodium acetate and ethanol. The cut DNA was resuspended in 25 µl of TE, 25 µl of 2X Bal 31 buffer (24 mM $CaCl_2$, 24 mM $MgCl_2$, 40 mM Tris-HCl pH 8, 2.0 mM EDTA, and 0.8 M NaCl) was added, and the tube was stored at 30°C. Bal 31 Stop solution (25 mM EGTA pH 8) was aliquoted into 16 labelled tubes (15 µl/tube). 1 unit of Bal 31 enzyme (New England Biolabs) was added to the DNA (at 30°C), and 3 µl aliquots were removed from the reaction tube every 5 minutes, added to a tube of stop solution, and stored on ice. When the Bal 31 digest was completed, 52 µl of water, 8 µl of restriction enzyme buffer, and 20 units of *Eco* RI was added to each of the 16 tubes and they were incubated at 37°C for 1 hour. The tubes were heated to 70°C for 15 minutes to inactivate the *Eco* RI, and then an 18 µl aliquot from each tube was run on a 3.5% acrylamide gel with appropriate markers. The remaining digested DNA was stored at -20°C. Inspection of the gel was used to determine which tubes contained DNA fragments digested to the appropriate lengths. 5 µl of DNA from each of these tubes was ligated to 100 ng of the 'cassette' M13 vector described above (reaction volume was 25 µl), and the ligation mixture used directly to transform competent JM 109 cells. Although there was a mixture of two DNA fragments present in the Bal 31/*Eco* RI digests, both of which had one blunt and one *Eco* RI 'sticky' end, the vast majority of the M13 transformants contained the shorter cDNA fragment rather that the fragment consisting of the plasmid vector. That the M13 phage transformants contained a portion

of the cDNA insert was simply confirmed by performing a C-test with a phage containing the full length cDNA insert in the opposite orientation. This method of generating deletions has the advantage that phenol-chloroform extraction of all of the tubes from the various time points is not required (as is the case with other exonuclease digestion procedures), since Bal 31 can be irreversibly inactivated with EGTA (Gray *et al.*, 1975). Furthermore, by simply inspecting the gel the exonuclease digestion can be monitored and modified as necessary (*e.g.* by increasing the time of incubation or amount of enzyme, or by decreasing the salt concentration).

b) Preparation of Template DNA. Isolated M13 'plaques' were picked with the sharp end of a sterile pasteur pipet, placed into culture tubes containing 2 ml of 2X YT media (Amersham Sequencing Handbook) and grown overnight with shaking at 37°C. The bacteria were removed by centrifugation (5 minutes in the microfuge), and supernatants were removed and stored at 4°C. These represent permanent phage stocks which remain viable for years. To grow templates, 2 ml of a fresh overnight culture of JM 109 (grown in M9 media, and picked from an isolated colony on an M9 plate) was added to 50 ml of 2X YT, and this suspension was aliquoted into the required number of culture tubes (2 ml/tube). 30 µl of phage stock was added to each tube (usually 2 tubes were innoculated with each stock) and the phage was grown with vigorous shaking (≥ 325 rpm on the air shaker; aeration is the single most important step in preparing template) for exactly 6 hours. The bacteria were pelleted (5 minutes in a microfuge) and the supernatant transferred to a second tube (keeping well away from the pellet). The second tube was centrifuged for a further 5 minutes, and supernatants transferred to a third tube, again keeping well away from the pellet. (The second centrifugation ensures that no bacterial debris remains in the supernatant.) Phage particles were precipitated by adding 300 µl of 20% polyethylene glycol (PEG), 2.5 M NaCl (in SM buffer; Maniatis *et al.* 1982) and storing the tube on ice for 20 minutes.

The phage was collected by spinning 7 minutes in the microfuge, and the supernatant removed and discarded. The last traces of supernatant (and PEG) were removed by centrifuging the tube for a futher 3 minutes and using a drawn-out pasteur pipet to remove the remaining drops of liquid. The visible phage pellets were resuspended in 350 µl of TE, and solutions of identical phage were pooled (total volume was usually 700 µl/tube). The phage solutions were extracted 2X with 400 µl phenol (equilibriated with TE), 2X with 400 µl phenol-chloroform (1:1, equilibriated with TE), and 2X with 400 µl chloroform. DNA template in the final aqueous phase was precipitated by adding sodium acetate to 0.3 M, filling the tube with ethanol, and storing it at -70°C for several hours. DNA pellets were rinsed with 70% ethanol, dried, and resuspended in 50 µl of water. The quality and purity of the template was checked by running a 4 µl aliquot on a 0.7% agarose gel.

c) The Sequencing Reaction. The sequencing reactions were done on terizaki microtitre plates (Falcon Plastics), using a variation of an unpublished 'petri plate sequencing' protocol developed by G. Pielak (U.B.C.). To anneal the templates to the primer, 5µl of template (prepared as described above), 1.5 µl of annealing buffer (100 mM Tris-HCl pH 8.5, 100 mM $MgCl_2$), 1 µl of primer (17 bp, 0.03 OD_{260} per ml), and 3 µl of water were combined in a 1.5 ml Eppendorf tube and incubated for 1-2 hours in an oven at 60°C. The terasaki plate was labelled (rows G,A,T,C; columns according to templates), leaving the top row of wells unlabelled. The plate was placed on ice, and to each of the wells in the top row was added the following: 1 µl klenow fragment (5U/µl, Promega Biotec), 1 µl DTT (0.1 M), 1 µl dATP (13µM), and 2 µl [alpha-^{35}S]-dATP (>1000Ci/mmol, New England Nuclear). 2 µl of 'G-mix' was added to each of the wells in the row labelled G, and the same was done for the rest of the rows using the corresponding reaction mixes. The composition of the reaction mixes is given in Table 2. The template with annealed primer was transferred to the

TABLE 2

PREPARATION OF DIDEOXY REACTION MIXES FOR DNA SEQUENCING

1. Deoxy NTP mixes* (dA*, dC*, dG* and T*)

	dA*	dC*	dG*	T*
0.5 mM dCTP	200µl	10µl	200µl	200µl
0.5 mM dGTP	200µl	200µl	10µl	200µl
0.5 mM TTP	200µl	200µl	200µl	10µl
TE buffer	200µl	200µl	200µl	200µl

2. Dideoxy NTP mixes (ddC, ddG, ddT, ddA)

	ddC	ddG	ddT	ddA
5.0 mM ddNTP	1µl	1µl	3µl	1µl
Water#	240µl	120µl	120µl	240µl

3. Reaction mixes: prepared by adding equal volumes of the deoxy mixes and the dideoxy mixes; e.g. 100µl dN* + 100µl ddNTP

*from the Amersham M13 cloning and sequencing handbook

#the amount of water may be varied by a factor of two as necessary to obtain a clear sequence over the desired range

appropriate well in the first row, and mixed by pipetting up and down several times. 3 µl of the resulting mixture was then transferred to each of the four wells (G,A,T,C) immediately below it, taking care to mix each well by pipetting, and changing pipet tips between each transfer. The entire plate was incubated in a shallow pan of water in the 37°C incubator for 20 minutes. 2 µl of 'chase' mix (0.5 mM of all 4 dNTPs) was added to each well and the reaction incubated a further 20 minutes. The reaction was terminated by adding 4 µl of formamide dye to each well, the products were denatured by placing the plate on a water-filled heating block at ≈95°C for 5 minutes, and 1-2 µl was loaded on each track of a sequencing gel. The loading was facilitated by using a 10 µl Hamilton syringe with a 26 guage needle. (The outer diameter of the needle had been reduced by placing it in a drill and sanding the outer surface with emery cloth.)

d) Sequencing Gels. Buffer gradient gels were used for all DNA sequence determination. These gels were poured essentially as described by Biggin *et al.* (1983), except that sucrose was not added to the 5.0X TBE acrylamide mix, and 7.0 ml of the 5.0X TBE mix and 6.0 ml of the 0.5X TBE mix were used to construct the initial gradient. The glass plates used to form the gels measured 50 x 20 cm, and 0.4 mm thick spacers were used (BRL). A 20 x 40 cm aluminum sheet (6061) was placed in contact with the anterior glass plate to prevent 'smiling.' 0.5X TBE was used in the upper buffer chamber, and 5.0X TBE in the lower chamber, and these concentrations reflect the gradient of salt present in the gel itself. A 'mini-shark's tooth' comb (BRL) was used to form a total of 48 separate wells per 20 cm gel (each well 3 mm wide), and therefore a total of 12 sequences could be run on a single gel. The gels were run at 40 watts constant power for a total of 3.5-4.0 hours, and this meant that 250-300 bases of sequence could usually be read for each template loaded. The gels were fixed in 10% methanol, 10% acetic acid for 15 minutes, covered in saran wrap, and dried for 1-2 hours at 80°C in a gel dryer (Biorad). The saran wrap was removed and the gels

exposed overnight at room temperature (XRP1 film). DNA sequences were read and recorded manually. The sequence data was stored and analyzed using the program 'DNA Inspector II' (Textco) on a Macintosh 512K computer (Apple Computers). Searches of DNA and protein data bases were done on a variety of systems (Bionet, Microgenie, Gene Pro) as convenient.

3. *Expression of PMA-Induced cDNA Clones in COS Cells*

The final step in the characterization of those cDNA clones known or suspected to code for lymphokines was to insert them into a SV40 based expression system and in turn transfect these constructs into COS cells. In this way the proteins encoded by the various cDNAs were produced and could then be tested in relevant bioassays. The expression vector pcEXV-3 was obtained from J. Miller (NIH, Bethesda). In this plasmid, transcription of inserted cDNA is driven off the SV40 early promoter, and SV40 poly A addition signals are included downstream of the cDNA (Miller *et al.*, 1985). The cDNA can be cloned into either a *Sma* I or *Eco* RI site. The plasmid was cut with *Eco* RI and dephosphorylated, purified cDNA inserts were ligated into the cut site, and recombinants were obtained with the insert in either orientation (as determined by restriction mapping). Plasmids with the insert in the 'backwards' orientation served as negative controls for transfection and expression. Transfection of COS 7 cells was accomplished using DEAE-dextran (Sigma #D-9885) and a glycerol shock, as described by Lopata *et al.* (1984). Briefly, for each transfection 1 x 10^6 COS 7 cells were seeded in a small plastic T-flask. The following day the monolayers (now confluent) were washed twice with RH and drained. A mixture was made by adding 5-10 µg of plasmid DNA ('miniprep' or gradient purified, sterilized by ethanol precipitation) to 1.5 ml of RH, and then adding 0:5 ml of a DEAE-dextran stock solution (2.0 mg/ml in RH, filter sterilized). (When added directly to the

DEAE-dextran solution the DNA will precipitate.) The DEAE-dextran DNA mixture was placed on the rinsed monolayer, and the flask was incubated at 37°C for 4 hours. The DEAE-D/DNA mix was removed by suction, 5 ml of 20% glycerol (in RH, filter sterilized) was added to the flask, and the monolayer was incubated for 2.5 minutes at room temperature. The glycerol was removed by suction, the flask was rinsed twice with RH, and 5 ml of RHFM was added. After 72 hours in the incubator culture supernatants were harvested, centrifuged to remove cells and debris, and stored at -20°C. The COS 'sups' were assayed directly for biological activity. Where material was sent to other cities it was lyophilized and then reconstituted and filter sterilized on arrival.

H. CAP LABELLING AND HYBRIDIZATION TO MMTV SEQUENCES

Poly A^+ RNA (from untreated EL4, or cells cultured with PMA or PMA+CSA) was 'de-capped' as described previously (Moss and Koczot, 1976). Briefly, periodate oxidation of 40 µg of RNA was done in a 0.11 ml reaction volume containing 0.15 M sodium acetate pH 5.3 and 0.9 mM sodium periodate; the reaction was incubated at 0°C for 30 minutes in the dark, then RNA was precipitated twice with ethanol and sodium acetate. ß-elimination was accomplished by redissolving the rinsed, dried pellet in 0.100 ml of a freshly prepared solution of 0.300 M aniline and 0.010 M acetic acid (adjusted to pH 5 with concentrated HCl) and incubating at room temperature for 4 hours. The RNA was again ethanol precipitated through two cycles and redissolved in 50 µl water. Recoveries varied from 25-50% of starting RNA. Cap labelling and hybridization were done essentially as described (Wittek *et al.*, 1980). Briefly, 4 µg of decapped RNA was incubated for 60 minutes at 37°C in a 0.1 ml reaction containing 50 mM Tris-HCl pH 7.9, 1.25 mM magnesium chloride, 6 mM potassium chloride, 2.5 mM DTT, 60 µM S-adenosylmethionine, 160 µCi of [alpha-^{32}P]-GTP

(>3000Ci/mmol, NEN), and 10 units of guanylyltransferase (BRL). The reaction was terminated by addition of EDTA to 10 mM, and labelled RNA was separated from unincorporated counts by sephadex G-50F gel filtration. Approximately $5x10^5$ cpm of cap labelled RNA was hybridized to each 14cm x 4cm nylon filter in 2.5 ml of 6X SSC for 24 hours at 60°C. Filters were washed in 4X SSC at 37°C, incubated with 25 µg/ml RNase A in 2X SSC at 42°C for 90 minutes, then washed thoroughly in 4X SSC and exposed at room temperature.

CHAPTER III

INDUCTION OF INTERLEUKIN 2 MESSENGER RNA INHIBITED BY CYCLOSPORIN A

A. INTRODUCTION

Cyclosporin A (CsA) is an effective immunosuppressive drug which has proven useful in organ transplantation. Its mode of suppression may be based, at least in part, on its ability to inhibit the production of the lymphokine Interleukin 2 (IL2) from T lymphocytes (Britton and Palacios, 1982; Bunges *et al.*, 1981; Reem *et al.*, 1983). In contrast, stimulation of T lymphocyte growth by IL2 is not sensitive to CsA (Orosz *et al.*, 1982; Larsson, 1980; Bunges *et al.*, 1981). Failure to detect IL2 synthesis in leukocyte populations in the presence of CsA, however, is not definitive evidence that IL2 is not being generated. Rapid utilization or destruction of the secreted lymphokine could account for its absence. Furthermore, even if it can be shown that IL2 is not being produced, it is interesting to determine if the block created by CsA is at the level of transcription, translation, or post-translational processing. In this context it would be useful to determine the effects of CsA on the levels of IL2 mRNA.

The mouse T lymphoma cell line EL4 (Farrar *et al.*, 1980) produces a number of lymphokines in response to phorbol-12-myristate-13-acetate (PMA), and the human T leukemia cell line Jurkat (Gillis and Watson, 1980) synthesizes IL2 on stimulation with PMA plus mitogen. In this chapter we show that both EL4 and Jurkat lines are inhibited from producing IL2 by CsA, and that this inhibition is correlated with an absence of IL2 mRNA as measured biologically or chemically. Furthermore, CsA similarly inhibited the induction of mRNA encoding colony stimulating factor for granulocytes and macrophages (GM-CSF) in EL4 cells. However, the effect of CsA was not due to a simple inhibition of all PMA effects.

B. CsA BLOCKS THE PRODUCTION OF IL2

The subclone EL4.E1 of the cell line EL4 was selected for its high level of IL2 secretion upon stimulation with PMA. In comparison with EL4.E1, the subclone JEH.3B of the human Jurkat cell line generates somewhat lower levels of IL2 upon stimulation with PMA plus Con A. To determine whether the synthesis of IL2 by EL4.E1 and Jurkat JEH.3B cells was sensitive to CsA, both lines were stimulated in the presence of various concentrations of CsA. CsA blocked the secretion of IL2 by both cell lines (Figure 3). The concentration of CsA required for 50% inhibition, 5-10 ng/ml, was similiar to concentrations previously found to be effective (Bunges *et al.*, 1981). Nearly complete inhibition was observed at 50 ng/ml, and this concentration was used in subsequent experiments. As others have found (Britton and Palacios, 1982; Bunges *et al.*,1981), CsA had no effect on DNA synthesis in the IL2-dependent mouse T-cell line used for assaying IL2.

C. CsA TREATMENT RESULTS IN A LOSS OF TRANSLATABLE IL2 mRNA

To determine whether the lack of IL2 secretion in the presence of CsA was due to a lack of mRNA coding for this lymphokine, we isolated mRNA from cells and translated it in *Xenopus laevis* oocytes (Bleackley *et al.*, 1981; Bleackley *et al.*, 1983). The biologically active translation products were assayed on the cloned murine cytotoxic T-lymphocyte line MTL2.8.2 (Bleackley *et al.*, 1982). Stimulation of EL4.E1 cells with PMA induced generation of IL2 mRNA, as detected by translation in oocytes (Table 3). When CsA was also present in the cell cultures at 50 ng/ml, PMA did not lead to the accumulation of translatable IL2 mRNA. As a control, mRNA was prepared from cells stimulated in the presence or absence of CsA and mixed at harvest. The mixed mRNA preparation gave the expected yield of translated IL2, demonstrating

FIGURE 3. Effect of CsA on the production of IL2 by murine EL4.E1 cells. (A) EL4.E1 cells were grown in RHFM-10%FCS, collected, and suspended at a density of 2 x 10^6/ml in RHFM-2%FCS and PMA (40ng/ml). Portions (0.5 ml) of this suspension were added to an equal volume of medium containing 0 to 2000 ng of CsA/ml. Cells were cultured for 24 hours and the supernatants were collected by centrifugation and assayed for IL2 activity. A stock solution of CsA (1 mg/ml) (Sandoz lot OL 27-400N) was made up in dimethyl sulfoxide. Dimethyl sulfoxide at the concentrations added with CsA was found to have no effect either on IL2 production by EL4.E1 cells or on the IL2 assay itself. Culture supernatants were assayed at several dilutions on the murine CTL line MTL2.8.2 by monitoring uptake of [^{125}I]iododeoxyuridine into cellular DNA. Activity was obtained from the titration curve and expressed as ED_{30} units. The experiment was carried out in duplicate, with data from the two experiments indicated by circles and squares. (B) Jurkat JEH.3B cells were grown in the same manner as the EL4.E1 cells and resuspended (2 x 10^6 cells/ml) in RHFM-2%FCS, PMA (40 ng/ml), and Con A (66μg/ml). Portions were mixed with equal volumes of CsA-containing meduim and supernatants were harvested after 24 hours of incubation and assayed for IL2 on the MTL2.8.2 line [human IL2 is active on murine cells]. Con A and PMA at equivalent concentrations had no effect on the bioassay. The experiment was also carried out in duplicate.

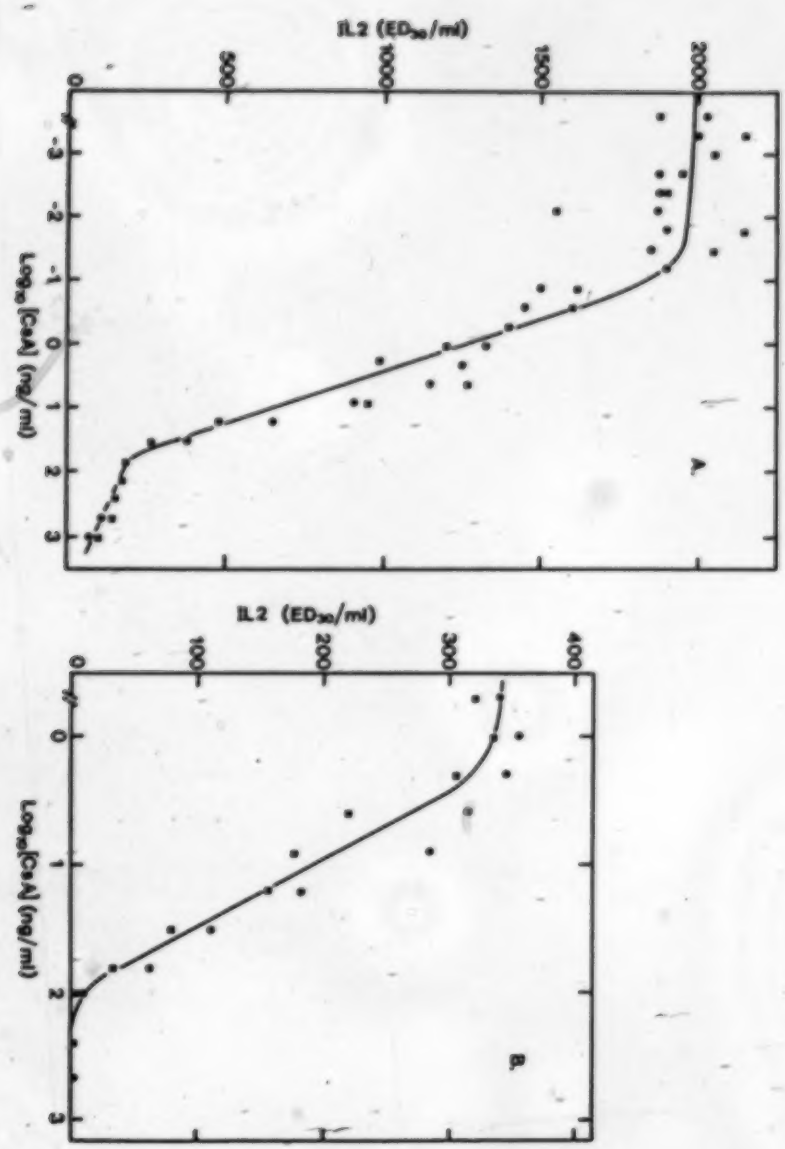

IL2 (ED50/ml)
0
500
1000
1500
2000
Log10[CsA] (ng/ml)
-3
-2
-1
0
1
2
3
A
IL2 (ED50/ml)
0
100
200
300
400
Log10[CsA] (ng/ml)
-1
0
1
2
3
B

TABLE 3

EFFECT OF CsA ON THE LEVELS OF TRANSLATABLE IL2 mRNA IN EL4.E1 CELLS

Poly A^+ RNA from EL4 cells cultured with:	IL2 activity CPM: oocyte supernatant	oocyte homogenate	Total (U/ml)
A. 20 ng/ml PMA	14,900 ± 240	9700 ± 450	128 ± 7
B. 20 ng/ml PMA+ 50 ng/ml CsA	113 ± 1	195 ± 19	1.4 ± 0.2
C. Medium only	55 ± 9	165 ± 11	0.4 ± 0.3
D. Equal numbers of cells from A+B	10,449 ± 139	605 ± 118	60.0 ± 1.1

EL4.E1 cells were grown in RHFM-10%FCS. Cells in log phase growth were collected and resuspended at 1 x 10^6/ml in RHFM-2%FCS. A 1.5 L culture was treated with PMA (A.), a second was treated with PMA plus CsA (B.), and 1 L was incubated without stimulation (C.). After 16 hr, cells were harvested and 3.5 x 10^8 cells from each of groups A. and B. were mixed to form group D. This group serves as a control to show that CsA-treated samples do not inactivate IL2 mRNA or prevent its translation. Groups A., B., and D. each contained about 7 x 10^8 cells (PMA stops proliferation, even in the presence of CsA). Group C., which had continued to proliferate, contained 1.8 x 10^9 cells. Cells were homogenized in 7.4 M guanidine HCl, and poly A+ RNA isolated. The mRNA was injected (≈50 nl at 2 mg/ml, 20 oocytes/group) into Xenopus laevis oocytes and these were incubated at 25°C for 63 hours. Media from each group were pooled ('oocyte supernatant'), and the oocytes were then pooled and homogenized ('oocyte homogenate'). Supernatants and homogenates were assayed on the mouse cytotoxic T lymphocyte line MTL2.8.2. The CPM shown were obtained at 1/20 dilution. Data from the activity-dilution curves were converted to total ED_{30} units/ml ± standard deviation recovered from each set of oocytes (last column).

that the lack of mRNA in the CsA-treated cells was not due to inhibitory material (Table 3). These results indicate that CsA reduced the level of IL2 mRNA by a factor of at least 10.

To confirm that lack of translation was not due to a peculiarity of the oocyte system, some mRNA samples were also translated in the cell free wheat germ extract system. This system is much less efficient than microinjected oocytes; however, the results (Table 4) lead to the same conclusion. No biologically active IL2 mRNA could be isolated from EL4.E1 cells stimulated in the presence of CsA.

Similar results were obtained with the line JEH.3B. Although the yield of IL2 and its mRNA was lower with this cell line than with EL4.E1, the effect of CsA was the same. When CsA was present during induction of IL2, no translatable IL2 mRNA was detected (Table 5). The mRNAs were also translated in wheat germ extract, with the same pattern of results as for EL4.E1 mRNA.

The translatability of various mRNAs used in these experiments, as assayed by the incorporation of [^{35}S]methionine into protein in wheat germ or rabbit reticulocyte lysates, was nearly the same. The general pattern of protein products synthesized, as analyzed by gel electrophoresis was also similar. Other experiments indicate that CsA is highly selective in its effect. Only a few (six to eight) proteins secreted by PMA-induced EL4 cells are sensitive to CsA (Ng, 1984), and analysis of cloned complementary DNAs indicates that $\leq 3\%$ of mRNAs expressed by EL4 are sensitive to CsA (see Chapter VI). Thus CsA does not affect the yield or general character of mRNA from these cell lines, and the reduction of IL2 mRNA levels is a specific effect.

D. NO DETECTABLE IL2 mRNA IN CsA TREATED CELLS

Sequences coding for human (Taniguchi *et al.*, 1983) and mouse (Fuse *et al.*, 1984) IL2 mRNA have been cloned. With cloned complementary DNA, or

TABLE 4

TRANSLATION OF IL2 mRNA IN WHEAT GERM EXTRACT

Source of mRNA added to translation mix	IL2 activity
No mRNA	78 ± 10
A. Unstimulated EL4.E1 cells	81 ± 22
B. EL4.E1 cells stimulated with PMA	543 ± 38
C. EL4.E1 cells stimulated with PMA + CsA	63 ± 6
D. Equal mixture of B.+ C.	384 ± 40

The same poly A^+ RNA used in Table 3 was added to a wheat germ translation mixture at 120 µg/ml. Samples were assayed for IL2 at 1/25 dilution, with the results being given as CPM of [^{125}I]-dIUrd incorporated into the assay cells.

TABLE 5

EFFECT OF CsA ON THE LEVELS OF TRANSLATABLE IL2 mRNA IN HUMAN JURKAT CELLS

Poly A^+ RNA from cells cultured with:	IL2 activity CPM: oocyte supernatant	oocyte homogenate	Total (U/ml)
A. 20 ng/ml PMA + 33 µg/ml Con A	315 ± 12	993 ± 19	5.49 ± 0.09
B. 20 ng/ml PMA + 33 µg/ml Con A + 50 ng/ml CsA	50 ± 8	158 ± 8	0.00 ± 0.29
C. Medium Only	31 ± 1	259 ± 63	0.48 ± 0.36

Jurkat JEH.3B cells were grown in RHFM-10%FCS, collected, and divided into 3 1L cultures at 1 x 10^6 /ml. Cultures were incubated with the additions indicated for 16 hours. Approximately 7.0 x 10^8 viable cells were recovered from groups A and B; 1.3 x 10^9 cells were recovered from group C. Poly A+ RNA was extracted from each group of cells and translated in oocytes as in Table 3. Oocyte supernatants and homogenates were assayed for IL2 activity at 1/20 dilution, and are given as CPM of [^{125}I]-dIUrd incorporated. Data from the activity-dilution curves were converted to total ED_{30} units ± standard deviation recovered from each set of oocytes (last column).

oligonucleotide probes derived from the published sequences, it is possible to measure IL2 mRNA by hybridization, independent of its translatability. Messenger RNA isolated from JEH.3B cells stimulated with PMA and Con A was separated by gel electrophoresis under denaturing conditions, the gel was blotted onto nitrocellulose, and the blot was probed with cloned human IL2 complementary DNA (Figure 4). A single band of the expected size (marker position, 12.5S) was detected in samples from stimulated cells. However, the corresponding band could not be detected in mRNA from cells stimulated in the presence of CsA. A virtually identical result was obtained when a similar northern blot was prepared using RNA from EL4 cells (Figure 5, right half), and probed with a 34 base oligonucleotide probe which hybridizes to murine IL2 mRNA (see Chapter IV). A single band of the expected size is seen in the RNA from PMA treated cells, but this band is absent in RNA from control cells or cells treated with PMA and CsA.

E. CsA ALSO BLOCKS THE mRNA FOR GM-CSF IN EL4 CELLS

Another lymphokine, GM-CSF, is also known to be secreted by EL4 cells when they are treated with PMA. Furthermore, when poly A^+ RNA from PMA-treated EL4 cells was microinjected into oocytes, biologically active GM-CSF was obtained (Bleackley *et al.*, 1983). The cDNA encoding murine GM-CSF has been cloned (Gough *et al.*, 1984), and from the published sequence several cross-hybridizing oligonucleotide probes were generated (see Chapter IV). A mixture of two to these probes were labelled and used to probe a northern blot made from RNA from EL4 cells (Figure 5, left half). A single band of the expected size is seen in the RNA from PMA-treated cells, but the corresponding transcript is absent in untreated EL4 cells or in cells treated with both PMA and CsA. This confirms that CsA also blocks the induction of GM-CSF mRNA in PMA-treated EL4 cells.

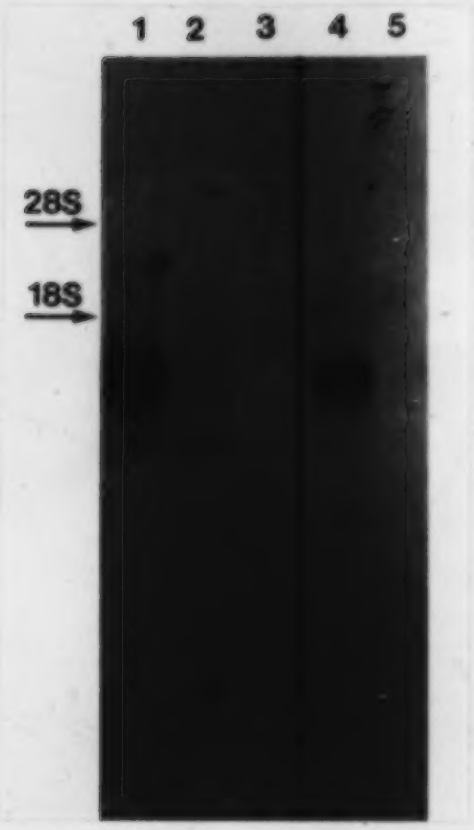


FIGURE 4. Absence of IL2 mRNA in CsA-treated cells. Polyadenylated RNA was isolated from JEH.3B cells, loaded onto a 1% agarose-formaldehyde gel, separated electrophoretically, and transferred to nitrocellulose. This was probed with a human IL2 cDNA probe, using standard conditions. The position of the 28S and 18S ribosomal RNA bands (detected on the original ethidium bromide stained gel) are indicated. Tracks 1 to 3 represent one experiment; tracks 4 and 5 another. Six micrograms of poly A+ RNA was loaded in each lane. Jurkat JEH.3B cells were treated as follows: tracks 1 and 4, PMA (20 ng/ml) and Con A (33 μg/ml); tracks 2 and 5, PMA (20 ng/ml) and Con A (33 μg/ml) in the presence of CsA (50 ng/ml); track 3, no stimulation. (Data from Yuan Lin; published in Elliott *et al.*, 1984)

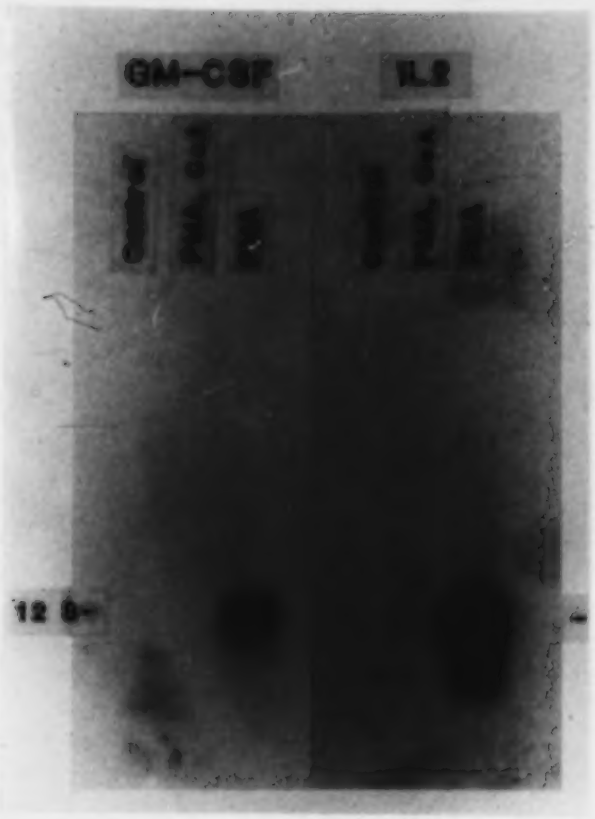


FIGURE 5. Induction of mRNA for IL2 and GM-CSF by PMA and supression by CsA in EL4.E1 cells. Cells were exposed to either PMA alone (20 ng/ml) or PMA (20 ng/ml) plus CsA (50 ng/ml) for 15 hours, and Poly A+ RNA was isolated. After separation by agarose gel electrophoresis, RNA was transferred to nitrocellulose and hybridized to synthetic oligonucleotide probes specific for either IL2 or GM-CSF (Table 6). Nine μg of poly A+ RNA was loaded per track. No mRNA for either lymphokine was evident in uninduced cells, or in cells induced in the presence of PMA plus CsA. (data from work done in collaboration with J. Shaw)

F. DISCUSSION

PMA-stimulation of EL4 cells results in the appearance of mRNA's coding for IL2 and GM-CSF, and the secretion of the corresponding proteins into the culture media. CsA blocks the induction of IL2 and GM-CSF secretion by EL4 cells, and the lack of these lymphokines correlates with the absence of their respective mRNAs. CsA similarly blocks the secretion of IL2 and the accumulation of IL2 mRNA in stimulated Jurkat cells. The applicability of these studies to normal human or murine lymphocytes must be confirmed by further experiments. It is important to note, however, that the effect of CsA on the induction of IL2 observed with these tumor cell lines occurred over the same range of CsA concentration (1-50 ng/ml) as was found to affect normal murine splenic lymphocytes stimulated with Con A (Bunges *et al.*, 1981). The CsA effect could be mediated by inhibition of IL2 mRNA transcription or by destabilization. It has been shown that insulin selectively inhibits synthesis of mRNA for the enzyme phosphoenolpyruvate carboxykinase in a cell line (Granner *et al.*, 1983); the present results would be explained if a similar inhibition of IL2 mRNA synthesis were caused by CsA.

Lymphokines in general and IL2 in particular play a central role in immune responses (Depper *et al.*, 1983). It appears that at least part of the immunosuppressive activity of CsA is due to its ability to selectively inhibit accumulation of mRNA coding for IL2 in human and murine T lymphocytes.

CHAPTER IV

ISOLATION OF cDNA CLONES ENCODING MOUSE IL2 AND MOUSE GM-CSF

A. INTRODUCTION

The objective of this study was to isolate and characterize a number of cDNA clones which contain sequences which are induced by PMA in the cell line EL4. An important positive control for the cDNA cloning process was to show that known PMA-induced, full length cDNA clones could be isolated from libraries generated using mRNA extracted from PMA-treated EL4 cells. Furthermore it would be useful to know the frequency of these clones in an unsubtracted cDNA library, so that this figure could be compared to that for a subtracted library which had been enriched for PMA-induced sequences. Such an unsubtracted library also serves as a 'back-up' if full-length cDNA clones cannot be found in the subtracted library, since the subtracted library is more likely to have shorter cDNA inserts, even if they have been size selected prior to ligation. The cDNAs isolated in these experiments would also be useful for a number of other experiments which are ongoing in Dr. Paetkau's laboratory.

At the time this work was done (September-December, 1984) oligonucleotide probes were available which cross-reacted with two of the lymphokine mRNAs known to be induced in PMA-treated EL4. These were murine IL2 and murine GM-CSF, and the structure of the oligonucleotides used is given in Table 6.

B. RESULTS

1. *Frequency and Size of cDNA Inserts in the Library*

Poly A^+ RNA (10 μg) from PMA-treated EL4 cells (harvested at 15 hours) was used to make 400 ng of size selected (>800 bp) ds-cDNA, as described in Chapter II. The cDNA was used to generate a library of ~5 x 10^6 *c* I⁻ phage. The frequency of clear

TABLE 6

DESCRIPTION OF SYNTHETIC OLIGONUCLEOTIDE PROBES

Probe	Length	Region encoded	Sequence
M-IL2-30	30	Amino acids 68-77 of primary translation product, murine IL2	(5') TTTGAAGGTGAGC-ATCCTGGGGAGTTTCAG
H-IL2-34	34	C terminus of human and murine IL2	(5')AGTTAGTGTTGAGAT-GATGCTTTGACAAAAGGTA
M-GM-43.13	43	Amino acids 76-89 of primary translation product murine GM-CSF	(5')GCTCGAATATCTTCAG-GCGGGTCTGCACACATGT-TAGCTTCTT
M-GM-43.14	43	Amino acids 90-104 of primary translation product murine GM-CSF	(5')TTCAAGGCGCCCTTGA-GTTTGGTGAAATTGCCCC-GTAGACCT

plaques (*i.e. c* I⁻ phage) in original ligation was ~0.27%, compared to a background frequency of ~0.13%, so that about 50% of the phage were expected to have cDNA inserts (see Chapter II). Five plaques from the library were picked at random, and DNA from each was prepared, cut with *Eco* RI, and run on a 1% agarose gel. Three of the phage had cDNA inserts, and two did not. All of the inserts were longer than 800 base pairs (one was 2200 bp, and the other two were ~1000 bp, data not shown).

2. *Isolation of Murine IL2*

Figure 6 gives an example of the autoradiograph which resulted when the 30 base IL2 oligonucleotide was used to probe a large plate containing ~25,000 independent recombinant phage clones (a total of 50,000 plaques per plate). The two signals circled are the only ones which also appeared when the duplicate filter was probed with the 34 base oligonucleotide probe. Figure 7 shows the autoradiograph of the rescreen of clone 17 from the previous figure, now probed with the 34 base probe. In total, 10 positive clones (which reacted with both probes) were obtained from 5 plates. Nine of the inserts were 800 bp in length, and the tenth (clone 17) was 960 bp. This 960 bp insert was subcloned into pUC 13. It was mapped with a number of restriction enzymes, and the map was shown to be identical to that of the published sequence of murine IL2 cDNA (data not shown).

3. *Isolation of Murine GM-CSF*

Two positive clones, which reacted with both probes for GM-CSF, were obtained from a second set of 5 large plates. Figure 8 shows an autoradiograph of a rescreen of one of these, probed with the 43 base '13' GM-CSF probe. Both clones had the identical sized cDNA insert (800 bp), and one of these (clone 3) was subcloned into pUC 13. It was mapped with a number of restriction enzymes, and the map was shown to be identical to that of the published sequence of murine GM-CSF cDNA (data not shown).

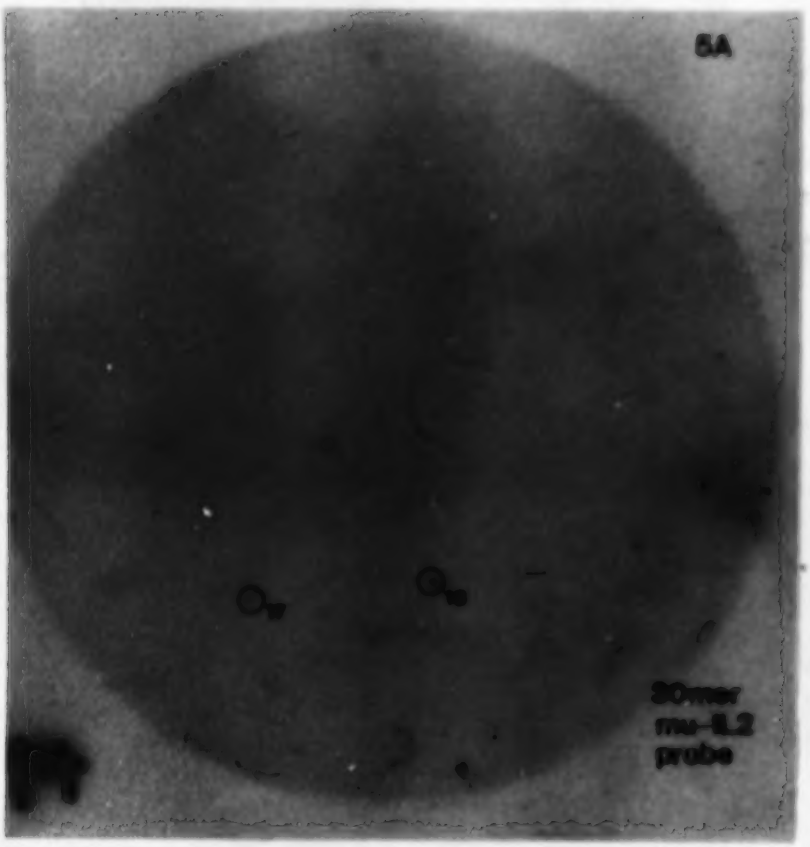


FIGURE 6. Unselected EL4 cDNA library probed with 30 base oligonucleotide probe homologous to mouse IL2. A 138 mm (diameter) nitrocellulose filter was used to make a replicate of a large petri plate which contained ~25,000 independent recombinant phage plaques. The filter was hybridized to the end-labeled probe, washed, and autoradiographed as described in Chapter II. The two positive signals which are circled also gave a positive signal on a duplicate filter probed with the 34 base probe (Table 6).

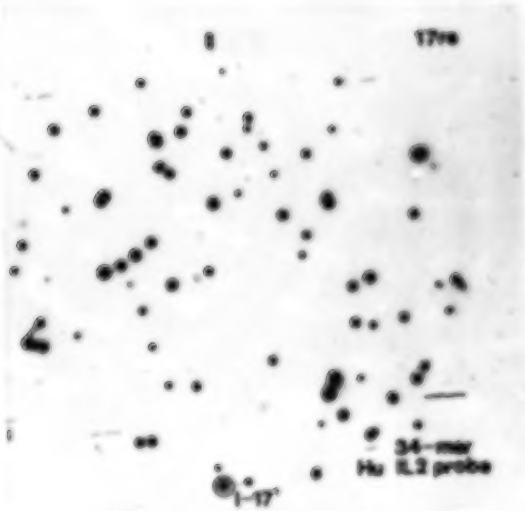


FIGURE 7. Rescreen of clone 17 using 34 base IL2 probe. The area of the plate indicated by 17 in the previous figure was 'picked' and the resulting mixture of phage was replated at a lower density (~750 plaques/90 mm petri plate). Most of the plaques do not give a positve signal, but number of positive plaques are seen. One of these was picked (1-17'; circled) and represents the final isolated recombinant.

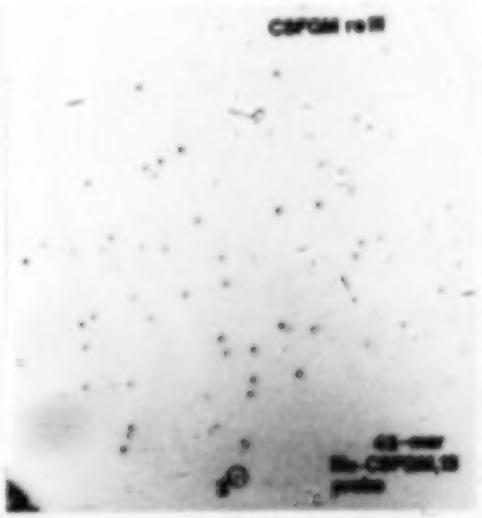


FIGURE 8. Final rescreen and plaque purification of a mouse GM-CSF cDNA clone (numbered CSF-GM 3'; circled).

C. DISCUSSION

There are two potential poly (A) addition signals within the mouse IL2 gene, and both of these appear to function to give rise to two possible transcripts with different 3' termini. The transcripts are about 800 and 930 bases (Fuse *et al.*, 1984), and the size of the cDNA inserts obtained in this study (800 bp and 960 bp) correlate well with the expected size of the mRNAs. It appears that the shorter transcript predominates in EL4.

If one assumes that IL2 is an mRNA of intermediate abundance (~300 molecules/cell), and that each cell contains ~300,000 mRNA molecules (Alberts *et al.*, 1983), then the expected frequency of IL2 in an 'ideal' cDNA library would be 1:1000. The apparent frequency of IL2 clones in the library screened in this study (1:12,500) is 12 fold lower than expected because the cDNA was stringently size selected prior to ligation into the vector. This means that the library is not a true representation of all of the mRNA's present in the cell, but is instead heavily biased in favor of cDNAs longer than ~800 bp. The majority of cDNA molecules copied from IL2 message are not likely to be full length, and therefore fall short of the 800 bp 'cut-off.' However, since the desired product is full length cDNA clones, the larger number of plaques which must be screened is more than compensated for by the fact that any clones found will contain all or nearly all of the original mRNA sequence.

The length of the predominant GM-CSF transcript is 790 bases (Miyatake *et al.*, 1985), and this compares favorably with the length of both GM-CSF cDNAs cloned (800 bp). The frequency of GM-CSF clones in this library was also much lower than expected for an 'ideal' cDNA library, and cannot be related to the actual frequency of the GM-CSF mRNA present in the cell for the reasons outlined above. However, the relative frequency of GM-CSF mRNA to that of IL2 mRNA can be estimated, since both transcripts are about the same length. The GM-CSF transcripts are about five times less abundant than the IL2 transcripts, since 2 GM-CSF clones were found,

compared to 10 IL2 clones, when the same number of plaques were screened. This agrees well with the relative abundance of the two transcripts as estimated by comparing the relative intensities of the two bands on northern analysis (Figure 5).

Each of the IL2 probes cross-hybridized to about 9 plaques per plate, although on average only two of these plaques gave a signal with both probes. The plaques that gave a signal with only one of the probes likely represent either artifacts or other cDNAs which have some sequence homology to the oligonucleotide probes. They probably do not represent incomplete IL2 cDNAs, since all of the inserts in the library were of sufficient length to make this very unlikely. The frequency of false positive clones found using a single ~30 base oligonucleotide probe makes it clear that two such probes should always be used when screening recombinant libraries. The longer, 43 base GM-CSF probes gave a lower frequency of false positive cross-reactive plaques (4-6 per plate), but again two probes (even from contiguous regions of the sequence, Table 6) made the task of selecting clones considerably easier.

This rather short chapter has presented data to demonstrate that the cloning method outlined in chapter II can be used to generate libraries which contain full length cDNA clones known to be induced in PMA-treated EL4 cells. The method used can be applied to the isolation of any cDNA clone, providing cross-reactive oligonucleotide probes are available. (In a separate set of experiments, the cDNA for human IL2 was isolated by screening a cDNA library generated from mitogen stimulated Jurkat cells.) The apparent frequency of the clones isolated in this study (1:12,500 for IL2, and 1:62,500 for GM-CSF) will be useful in estimating the enrichment factor for the subtracted cDNA library described in chapter 5, because similar size selection and cloning conditions were used to generate both libraries.

CHAPTER V

ISOLATION OF PMA-INDUCED cDNA CLONES FROM EL4 USING SUBTRACTIVE HYBRIDIZATION

A. INTRODUCTION

The object of this study was to isolate and characterize a number of PMA-induced cDNA clones from EL4. This chapter will consider the isolation of such clones, and the following chapter will describe their characteristics.

During the PMA-induced differentiation of EL4 cells, a limited number of new mRNAs are induced. This is in contrast to the vast majority of mRNAs, whose relative abundance remains unchanged during the PMA treatment (Paetkau *et al.*, 1984). The induced mRNAs may go from undetectable to detectable levels (*i.e.* a band on northern blots appears only in RNA from PMA-treated EL4), or they may simply increase over basal levels which are detectable in untreated EL4 cells. With any general strategy for isolating PMA-induced clones, cDNAs of both these classes will likely be discovered. Two cDNAs in the former class (IL2 and GM-CSF) were cloned as described in the previous chapter, and this provided an important positive control for the cloning process, as well as a test of the RNA to be used in later subtractive cloning.

A number of the strategies for cloning differential genes (as described in chapter I) can be applied to the problem of PMA-induced differentiation of EL4. Paetkau *et al.* (1984) used the 'plus-minus' method on a cDNA library cloned in pBR322 to isolate a cDNA clone ('probe 7') which increased about 4-fold after exposure to PMA. This clone has not been further characterized. Kwon and Weissman (1984) used differential or 'plus-minus' screening on a cDNA library from PMA treated EL4 (cloned in lambda phage BV-2). These authors tested the sensitivity of plaque hybridization for 'plus-minus' screening, using two reconstruction experiments, and concluded that this

method would detect clones which represented ≥ 0.02% of the total mRNA. They screened 30,000 recombinant phage, selected 200 potential induction-specific clones, and finished with 17 clones which consistently scored as inducible. On the basis of cross-hybridization these were placed into 6 groups, and members of the largest group (repesenting 50% of the clones) were sequenced and demonstrated to be derived from the *env* and 3'LTR region of mouse mammary tumor virus (MMTV). The remaining clones were not further characterized in their report. Apparently the clones selected did not include mouse IL2.

The 'plus-minus' method of screening for differentially expressed genes has a number of inherent limitations. A limited number of clones can be screened, since the plaques (or colonies) must be well isolated and easily identifiable on the plate, so that the signal from each probe can be compared on replicate filters. In practice therefore it is only possible to plate at a density of about 500-1000 plaques per 150 mm plate (St.John and Davis, 1979). The 'plus-minus' method also rests on the assumption that both of the duplicate filters have the same amount of DNA representing each plaque. The choice of positive clones is therefore rather subjective, and prone to artifacts, as demonstrated by the fact that of 200 potential PMA-induced cDNAs originally picked by Kwon and Weissman, only 17 (<10%) proved to be true positives. The process of screening using the 'plus-minus' method is limited by the background signal that all plaques (or colonies) give when probed with any cDNA, and this background can be the source of considerable difficulty, particularly when plasmid libraries (*i.e.* colonies) are screened. The screening and rescreening process also requires that the experimenter repeatedly compare large numbers of 'spots' on numerous filters, and this is tedious and exhausting, particularly if low abundance cDNAs (<0.01%) are sought. In fact the major shortcoming of any method which uses unsubtracted probe is that it results in only a low level of hybridization to sequences which are present at a low frequency.

For these reasons the author chose to use a strategy employing subtractive hybridization rather than 'plus-minus' screening to isolate PMA-induced cDNA clones from EL4 cells.

Subtractive hybridization was first used by Zimmermann *et al.* (1980) to isolate genomic clones which were induced in one developmental stage of *Aspergillus nidulans*. As discussed in chapter I, numerous other authors have used the technique to isolate a wide variety of differential genes. In particular, Hedrick *et al.* (1984a) used subtractive hybridization to isolate a cDNA clone encoding the ß-chain of the mouse T-cell antigen receptor. Most investigators have used subtractive hybridization to generate cDNA probe which is then used to screen random cDNA or genomic libraries. It has been estimated that screening random cDNA libraries for rare genes would involve surveying ≈1 million clones (Davis *et al.*, 1982). To overcome this problem, Hedrick *et al.* (1984a) generated a cDNA library which was highly enriched (50-fold) for sequences of interest by subtracting cDNA prior to the cloning process. Similarly generated subtracted probe was then used to screen a much smaller number of clones to isolate rare cDNAs of interest. In the work described below, a similar two stage approach was used by the author to isolate PMA-induced cDNA clones from EL4.

The isolation of differentially expressed genes using subtractive hybridization has a number of limitations. The hybridizations require (and consume) relatively large quantities of poly A+ RNA, and consequently large numbers of cells are required to generate the starting material. Generation of subtracted probe is technically demanding, and this is even more so for the case of constructing subtracted cDNA libraries. Relatively large quantities of radioactive nucleotides are required, and this can be expensive (but less so if the 'random primer' method of labelling subtracted probe is used; see chapter II). Compared to 'plus-minus' screening, the use of subtractive hybridization to isolate differential genes is more difficult and more expensive.

However, the latter technique has a number of advantages over the 'plus-minus' approach. Considerably larger numbers of clones can be screened, since plaques can be plated at a high density, and positives picked and plaque purified by rescreening. The technique is also probably faster and less tedious, since the experimenter does not need to compare a very large number of 'spots.' A question which has not been addressed is the reliability of the subtractive hybridization method for isolating differential genes. What fraction of the clones chosen during the initial screening represent induced sequences? The work described below will provide an answer to this question.

B. RESULTS AND DISCUSSION

1. *Subtracted Library*

Figure 9 gives an outline of the procedure used to generate the subtracted library, the details of which are described in chapter II. The original yield of ss-cDNA was about 70 µg, and following subtraction and gel exclusion chromatography about 7 µg (10% of the starting material) remained. If the apparent yield of single-stranded counts (5 million cpm) and counts eluted at 97°C (12.5 million cpm) are compared, the enrichment factor would appear to be about 4-fold. However, not all of the material loaded on the column can be eluted even at 97°C (the temperature at which the water in the column jacket boils), and presumably the residual material represents double stranded material, and sequences which have been removed from the starting cDNA. Furthermore, the material that flows through the hyroxyapatite column at 60°C is not all single stranded cDNA, but a combination of cDNA and degradation products, so that about 40% of the apparent single stranded counts are in fact 'free' ^{32}P counts or very short oligo-nucleotides which are not excluded from sephadex G50-F. Therefore the actual yield of the single stranded cDNA is ≈3 million cpm (0.6 x 5 million cpm) from a

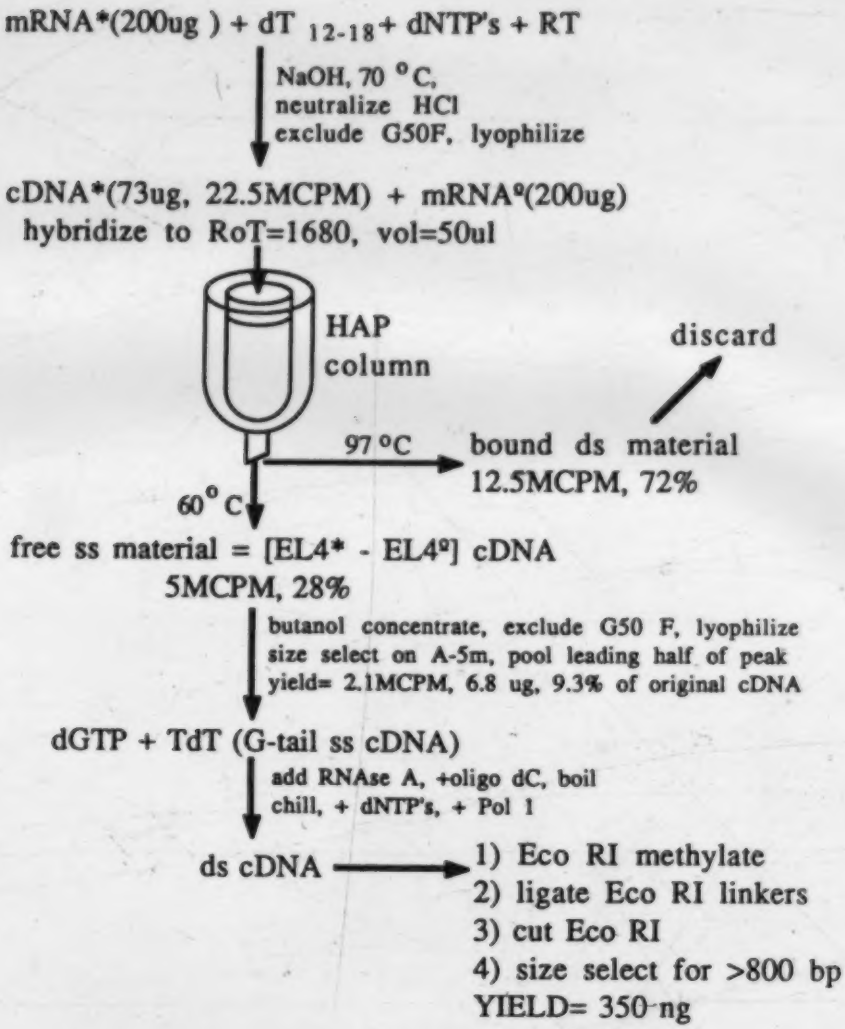

mRNA*(200ug) + dT 12-18 + dNTP's + RT
NaOH, 70 °C,
neutralize HCl
exclude G50F, lyophilize
cDNA*(73ug, 22.5MCPM) + mRNA°(200ug)
hybridize to RoT=1680, vol=50ul
HAP
column
discard
97 °C
bound ds material
12.5MCPM, 72%
60° C
free ss material = [EL4* - EL4°] cDNA
5MCPM, 28%
butanol concentrate, exclude G50 F, lyophilize
size select on A-5m, pool leading half of peak
yield= 2.1MCPM, 6.8 ug, 9.3% of original cDNA
dGTP + TdT (G-tail ss cDNA)
add RNAse A, +oligo dC, boil
chill, + dNTP's, + Pol 1
ds cDNA
1) Eco RI methylate
2) ligate Eco RI linkers
3) cut Eco RI
4) size select for >800 bp
YIELD= 350 ng

FIGURE 9. Procedure used to generate subtracted cDNA library. The yields after each of the various steps are given in micrograms of cDNA, million counts per minute of labelled cDNA (MCPM), and/or as a percentage of the total starting material. mRNA* represents poly A+ RNA from EL4 cells treated with PMA (20 ng/ml) for 14 hours [EL4*], and cDNA* is transcribed from this material. mRNA° represents poly A+ RNA from unstimulated EL4 cells [EL4°]. Complete details of the subtraction and cloning process are given in Chapter II. RT refers to reverse transcriptase; TdT refers to terminal transferase. To generate the amplified subtracted library 135 ng of the final size selected cDNA was ligated into lambda gt10 arms to yield one million independent recombinants. Only a small proportion of these were screened with the subtracted cDNA probe (see next figure).

starting amount of 22.5 million cpm, so that the estimated enrichment is about 8-fold (22.5/3 = 7.5). This level of enrichment was felt to be significant for the purposes of this study, and a second cycle of subtractive hybridization was not undertaken. Furthermore, each cycle of hybridization and hydroxyapatite chromatography appeared to result in significant degradation of the cDNA (considering the number of 'free' ^{32}P counts released), and it was felt that a second cycle of hybridization would result in cDNAs too short to be acceptable for cDNA cloning. In fact when this subtracted cDNA was run on a low melting point agarose gel (for the purposes of size selecting, see chapter II) it was seen to consist of a population of molecules which were considerably shorter (average size ≈480 bp) than those generated from the same cDNA which had not been subtracted (average size ≈700 bp) (data not shown). Although the subtractive hybridization process may have been selecting for cDNAs of a shorter size (*i.e.* the PMA-induced messages may consist largely of short mRNAs), it is also very likely that the hybridization and selection was causing considerable degradation of the cDNA. Although this effect has not been reported in the literature, it has apparently been observed by a number of investigators.

The ligation of the subtracted cDNA into lambda arms yielded a frequency of clear plaques (*i.e.* $c\ I^-$ phage) of 0.33%, compared to a background frequency of 0.16% in the control ligation where no cDNA was added. This means that about 50% of the phage in the cDNA library contained cDNA inserts. Nitrocellulose filters were replicated from four 150 mm plates (≈50,000 plaques/plate, or ≈25,000 recombinant phage/filter), and these filters were probed with oligonucleotide probes for IL2 and GM-CSF (the probes are described in chapter II). The frequency of IL2 clones (plaques which gave a signal with both IL2 probes) was ≈1:1400 recombinants, and GM-CSF clones occurred at ≈1:8000 recombinants. This compares to frequencies of 1:12,500 for IL2, and 1:62,500 for GM-CSF in the unsubtracted library described in

chapter IV. The apparent enrichment factor for IL2 clones (12,500/1400=8.9) and for GM-CSF clones (62,500/8000=7.8) in the subtracted library is in reasonable agreement with the estimated enrichment factor based on the amount of cDNA retained on the hydroxyapatite column (see above). The presence of these two known PMA-induced cDNAs at increased frequencies in the subtracted library made it very likely that the library contained, and was enriched for, other unknown PMA-induced sequences. To isolate such cDNAs the subtracted library was probed with subtracted cDNA probe.

2. *Subtracted cDNA Probe*

Figure 10 gives an outline of the procedure used to generate cDNA probe which was highly enriched for PMA-induced sequences in EL4. The quantities and yields are from a typical preparation, and in general were very similar each time subtracted probe was made. Two cycles of subtractive hybridization and hydroxyapatite chromatography were used, the first cycle resulting in approximately an 8-fold enrichment (912/109=8.4), and the second in a further ≈3-fold enrichment (109/36=3.0). The overall enrichment was 25-fold, so that about 4% of the starting cDNA remained. This represented a rather small quantity of probe (in terms of total number of counts), so that unusual measures were taken to hybridize in small volumes and at high temperatures (at which hybridization occurs about 3X faster than at 42°C), as described in chapter II.

As observed above, a relatively large number of 'free' ^{32}P counts or short oligonucleotides were released with each cycle of hybridization and subtraction (*e.g.* 173 million cpm appeared to be single stranded, but only 109 million cpm [63%] were excluded from a G50-F column during the first cycle). In the original 'cascade' hybridization procedure used by Timberlake (1980), Zimmermann *et al.*, (1980), and Mather *et al.* (1981) a final positive selection step was used to purify the cDNA of

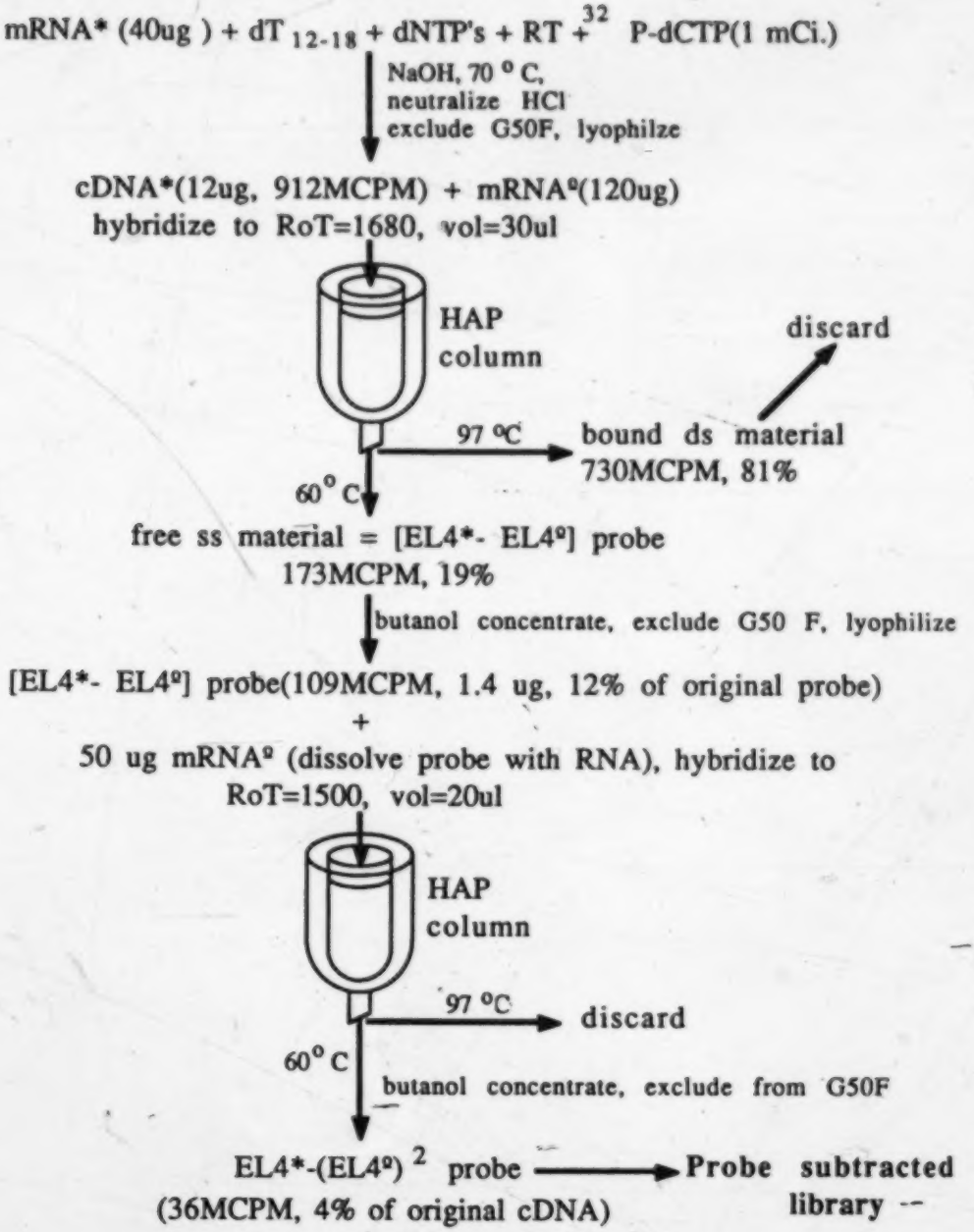


FIGURE 10. Procedure used to generate subtracted cDNA probe. The various preparations of mRNA and cDNA are as described in figure 9.

interest. The equivalent step in the series of experiments described here would be to hybridize the final subtracted probe to an excess of poly A^+ RNA from PMA-treated EL4 cells and then collect the double stranded material by a third cycle of hydroxyapatite chromatography. One of the effects of this step is to purify longer cDNA from 'free' ^{32}P counts and short oligonucleotides, a process which can be more simply and directly accomplished by G50-F gel exclusion chromatography. Such a 'back-reaction' also removes cDNA molecules which cannot hybridize to the mRNA used in the positive selection hybridization step. However, this 'unreactable' cDNA does not effect the screening process (either in terms of increasing the background signal or by giving false positive signals), and the 'back-reaction' is unnecessary to generate straight subtracted probe (M. Davis, personal communication).

The observation that only 4% of the original cDNA remained after two cycles of subtraction allows an estimate of the maximum difference between the mRNA populations produced by EL4 cells before and after PMA stimulation. A maximum of 4% of the total mRNA present in PMA-treated EL4 cells represents sequences which are induced, and the true figure is probably lower than this.

3. *Probing the Subtracted Library*

Figure 11 shows an example of an autoradiograph which resulted when the subtracted library was probed with subtracted cDNA probe. A large proportion of the positive signals are circled (in pencil). The original plate contained ≈50,000 *c* I⁻ phage, and about 25,000 recombinants which contained cDNA inserts. When probed with the subtracted probe ≈600 discrete positive signals could be detected per plate (average of four plates), so that about 1:42 (2.4%) of the recombinants gave a signal. Virtually all of the plaques giving a positive signal with the subtractive probe probably represent PMA induced cDNA sequences (see below). These results raise the following question, to which there is no particularly satisfactory answer. If ≈3% of the mass of

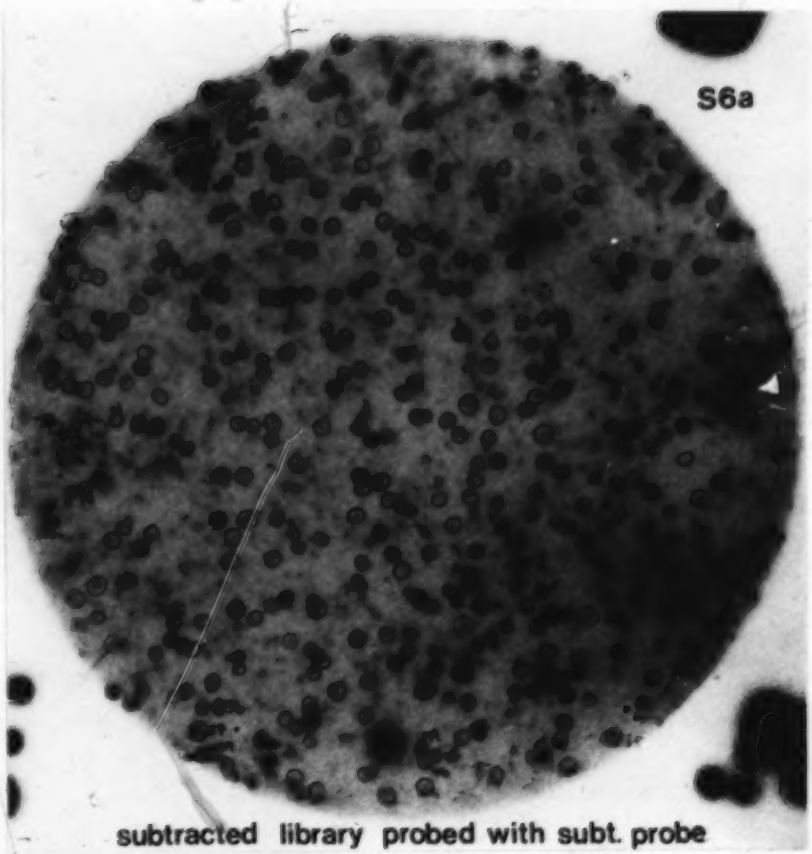


FIGURE 11. Subtracted cDNA library probed with subtracted cDNA probe. About 25,000 recombinant plaques were plated per 150 mm petri plate, and a number of replicate nitrocellulose filters were made from each of four plates. These were probed with subtracted probe, and a representative autoradiograph is shown. A number of the apparent positive plaques are circled in pencil, and those which are numbered were picked and plaque purified. [For example, clone number 25 is at 11 o'clock, clone 14 at 4 o'clock, and clone 12 at 7 o'clock; see below.] A replicate filter was subsequently screened with the IL2 probe, and areas which were positive are circled and labeled with a small dark 'dot' on the margin of the circle (there were on average 18 IL2 positive plaques per plate). The black marks at 3 o'clock represent an artifact which was present on the negative.

the mRNA in PMA-treated EL4 cells is induced, and if the subtracted library has enriched 8-fold for PMA induced sequences, then it is expected that ~24% (8 x 3%), or 1:4 recombinants should contain a PMA-induced cDNA sequence. However, the apparent frequency of such sequences is about 10-fold lower than expected. This is analogous to the observation made in chapter IV that the observed frequency of IL2 positive clones in an (unamplified) unsubtracted library was 12-fold lower than expected. In both libraries the cDNA inserts have been rather stringently size-selected, and apparently the resulting libraries do not give a true representation of all of the mRNAs present in the cell, but are heavily biased in favor of longer mRNAs. It is possible that not only IL2, but most of the PMA-induced messages are of a size that the majority of their cDNAs fall below the 800 bp 'cut-off.' Despite this apparent drawback, size selection is necessary to generate useful cDNA libraries which have long or full length cDNA inserts.

An alternate explanation for the discrepancy between the estimated and observed frequencies of PMA induced clones is that although ~50,000 plaque forming units are placed on each plate, because of the high density only 10% of these actually form plaques large enough to give a signal. This explanation is probably not correct, however, since if the subtracted library is plated at a much lower density (~100 plaques/plate), still only 2-3% of the plaques give a signal with the subtracted probe (data not shown). Whatever the reason for the discrepancy between the estimated and observed frequencies of PMA induced clones, subtractive probe still allowed the identification and isolation of sequences of interest, a task which would otherwise be extremely difficult if not impossible.

Although the frequency of a cloned sequence in the size selected, subtracted library cannot give an estimate of the absolute frequency of the given mRNA in the cell, the relative frequencies of mRNAs can be compared in this way if the mRNAs are of

the same size class (as was done in chapter IV for IL2 and GM-CSF), and if both mRNAs are induced by PMA so that both have been enriched the same amount by subtraction.

Figure 12 shows the autoradiograph resulting from the rescreen of one of the positive plaques (clone 25) which was originally picked from the subtracted library. In total, 136 independent clones (which gave a signal with the subtracted probe) were picked from the subtracted library, replated at an appropriate density, and rescreened (using a second preparation of subtracted probe, also generated as outlined in Figure 10) until they were plaque purified. The original plaques were numbered, and the numbers (for some of the clones) can be seen on the corresponding positive signals circled in Figure 11. In general the clones were picked from areas close to the 'alignment marks' on the plate, as described in chapter II. All clones that gave a signal on the original plate also gave a signal on the rescreen, demonstrating that the subtracted probe (prepared on two separate occasions) was reliably detecting certain cDNA sequences.

4. *Grouping the cDNA Clones into Cross-hybridizing Families*

A method for quickly grouping the 136 clones isolated above into a series of cross-hybridizing families is described in detail in chapter II. The cDNA inserts from a small number of clones (first subcloned into plasmids) were each used to probe a set of filters containing an ordered array representing the entire 136 clones. Figure 13 shows the autoradiograph resulting from two such experiments. In the top half of the figure, clone 25 was used as a probe. Only a single member of the 136 clones was capable of cross-reacting with the probe, and this was clone 25 itself (circled). Apparently this sequence occurred only once in the 136 clones, making it likely that it represents a lower abundance mRNA. In the lower half of the figure, clone 14 was used as a probe. Ninety-nine of the 136 clones cross-reacted with this probe, and this defined a very

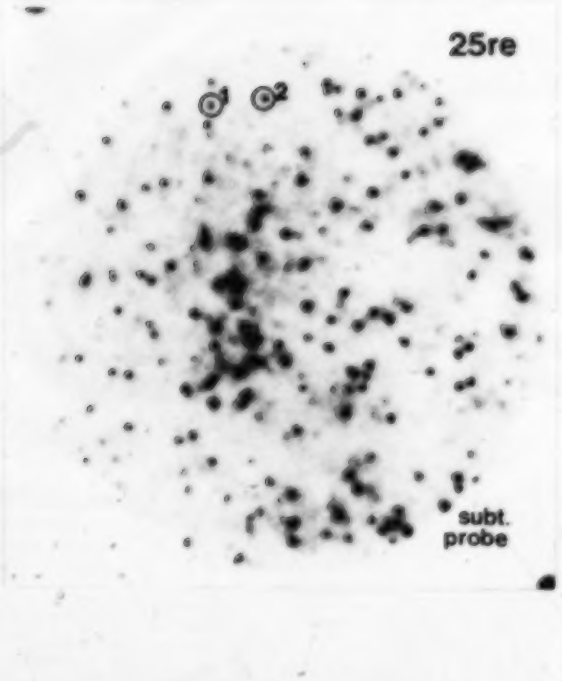
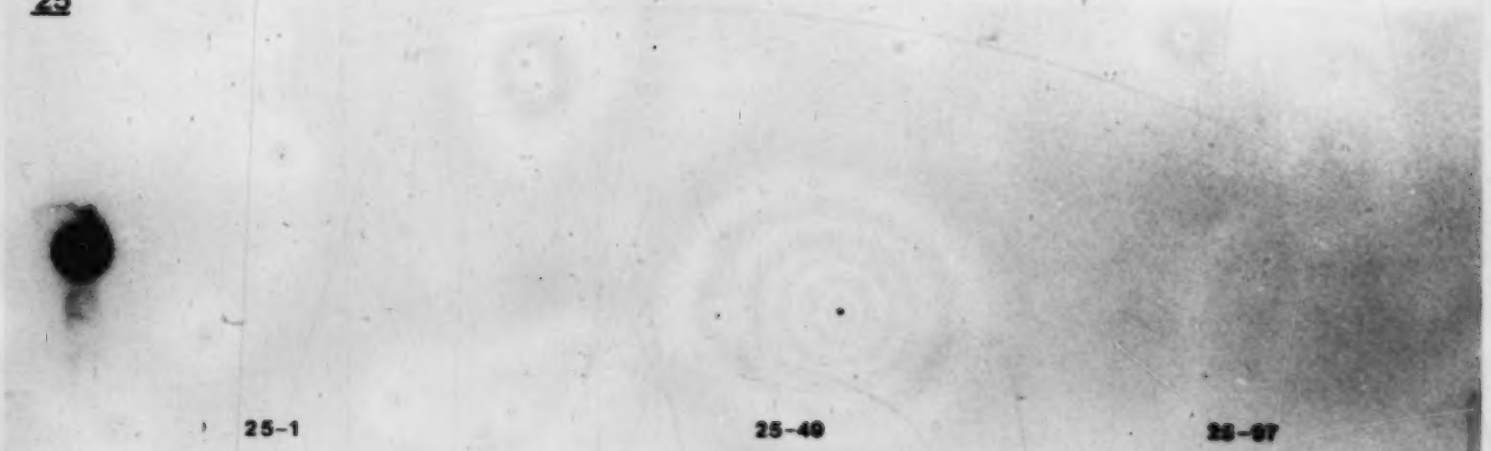


FIGURE 12. Rescreen and plaque purification of clone 25 cDNA. The region marked '25' in the previous figure was picked, and the resulting mixture of clones was plated at a lower density (~400 plaques/ 90mm petri plate). These were screened with a second preparation of subtracted cDNA probe, and although a large number of the clones gave a weak positive signal (contributing to a generalized background), a subset of the plaques gave a definite positive signal. Two such plaques (which were well isolated from surrounding plaques) were picked and represent isolates of clone 25 (circled and numbered 1 and 2). The cDNA insert from phage clone 25-1 was isolated and subcloned into a plasmid (pUC) vector.

FIGURE 13. Placing the PMA-induced cDNA clones into groups on the basis of cross-hybridization. The 136 phage clones which were plaque purified were innoculated onto fresh bactrial lawns in an ordered array, with 48 different clones per 90 mm petri plate. The phage plaques were allowed to grow up and then replicate nitrocellulose filters were made from each of the plates. For example, the filter in the bottom left (14-1) contains DNA from clones 1 to 48 (placed left to right, top to bottom), and the filter beside it (14-49) contains DNA from clones 49-96. The three filters in the top half of the figure were probed with labelled cDNA insert from clone 25, and the only clone which gives a signal is 25 itself (circled). The three filters in the bottom half of the figure were probed with clone 14 cDNA insert, and a large number of clones cross-react with this probe (clone 14 is circled). Beyond number 115 the clones are not numbered consecutively and this explains why it is possible to have a clone number 154 when only 136 clones were isolated (see below).

25



25-1 25-49 25-97

14

14-1 14-49 14-97

large family of clones (identical or closely related) which represent the most abundant PMA-induced transcript in EL4. Clones which did not cross-react with the initial set of probes were themselves subcloned and used to probe a similar set of filters. In this way all of the original 136 clones were placed into families and representative members of each were available in plasmids for further analysis. The relative number of clones in each of the various groups gives an estimate of the relative abundance of the corresponding mRNA in the total pool of all PMA-induced transcripts. A convenient summary of the different individual PMA induced cDNA clones found in this study is given in Table 7 (Chapter VI).

5. *Northern Analysis of Selected cDNAs*

Representative members of the families of cDNA clones selected and grouped as described above were subcloned into plasmids, and the cDNA inserts from these were labelled and used to probe northern blots of poly A^+ RNA from PMA-treated and untreated EL4 cells (the RNA used in these blots was in fact the same RNA used in generating both the subtracted library and the subtracted probe). This critical experiment was necessary to confirm that the cDNA sequences selected did indeed represent PMA-induced transcripts in EL4. The results of the northern analysis also give some indication of the size and nature of the transcripts which hybridize to the various clones.

Figure 14 demonstrates the results from one such northern, probed with clone 55. There is a single band (corresponding to a single transcript) which is induced only in PMA treated cells (lane p]), and absent in untreated EL4 cells (lane o]). Similar results were obtained for clones 12 and 25, although the sizes of the transcripts detected varies among the different clones (see Figures 19 and 21). Figure 15 demonstrates the northern results for clone 154; the autoradiograph shows a faint band present in lane o), which is more evident on a longer exposure (right half of Figure 15). This clone

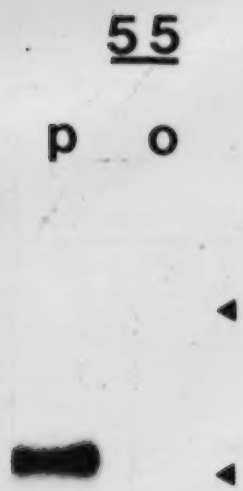


FIGURE 14. Clone 55 cDNA detects a PMA induced transcript in EL4.E1 cells. Polyadenylated RNA was isolated from EL4.E1 cells, separated by electrophoresis on a 1% agarose-formaldehyde gel, and transferred to nitrocellulose. This blot was probed with labelled cDNA insert from clone 55 using standard methods. Five micrograms of poly A+ RNA was laoded in each lane. Lane p), RNA from EL4.E1 cells cultured for 14 hours in the presence of PMA (20 ng/ml); lane o), RNA from unstimulated EL4.E1. The two samples of RNA are from the same stock used to generate the subtracted cDNA library and the subtracted cDNA probe. The arrows indicate the position of the 18S and 28S mitochodrial RNA bands as determined by ethidium bromide staining of the original gel.

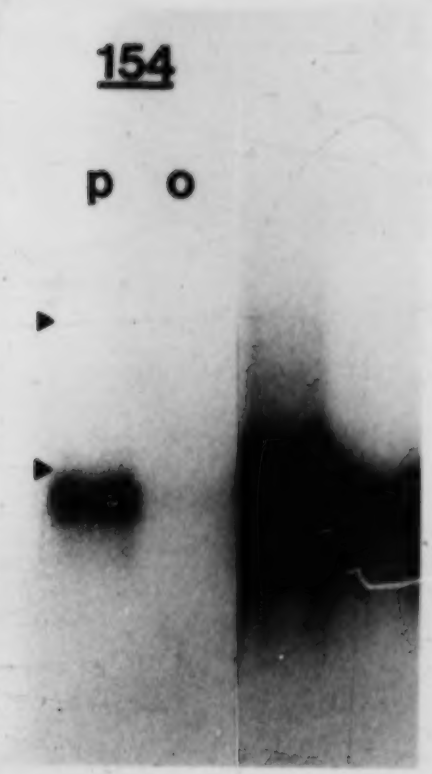


FIGURE 15. Clone 154 cDNA detects a PMA induced transcript in EL4.E1 cells. The two panels show the same northern blot exposed for different times. On the left, a short exposure shows that clone 154 detects a single band in lane p). However, a very faint band of the same size can also be seen in lane o), and this is more evident on a longer exposure, shown on the right. The northern blot used in this experiment is identical to the one described in Figure 14.

represents a sequence which is present at very low levels in resting EL4 cells, but induced to high levels following PMA induced differentiation. As discussed in the introduction, this is a class of transcripts one would expect to find using the subtractive strategy described. Figure 16 shows the northern results for clone 11. This sequence cross-reacts with a large number of transcripts, all of which are present at low levels in resting EL4 cells, and induced to high levels by PMA.

Figure 17 shows the results of three independent northern blots probed with clones 22, 15, and 14 respectively. All of these clones appear to cross-react with the same transcripts to produce nearly identical northerns. This result made it very likely that the three clones were derived from the same or a closely related mRNA, and this was indeed demonstrated to be the case (see lower half of Figure 13; clone 14 cross-hybridizes with clones 15 and 22, as well as many others). All of these clones cross-react with two larger transcripts which are present in resting EL4 cell and only slightly induced by PMA. However, they also detect a short transcript which is only present in PMA treated cells (best seen in the blot probed with clone 22). The final nothern blot in this series (Figure 18) was probed with clone 32. Although detected by the subtracted probe, this clone appears to be a false positive, since it is present at nearly equivalent levels in both resting and PMA-treated EL4 cells. The northern blot itself however provides an important control, since it demonstrates that nearly equivalent amounts of intact RNA are present in both the p) and o) lanes in all of the northern blots shown in this chapter (all the northerns used were transferred from the same denaturing gel, with equivalent amounts of RNA loaded in each track).

Collectively the northern results demonstrate that the clones selected using subtracted cDNA probe in almost all cases represent cDNA sequences which are induced by PMA in EL4 cells. One of the clones selected (clone 12) cross-hybridized with mouse IL2, and mapping and sequencing confirmed that it was IL2 (data not

FIGURE 16. Clone 11 cDNA detects a very large number of PMA induced transcripts in EL4.E1 cells. The northern blot used in this experiment is identical to the one described in Figure 14.

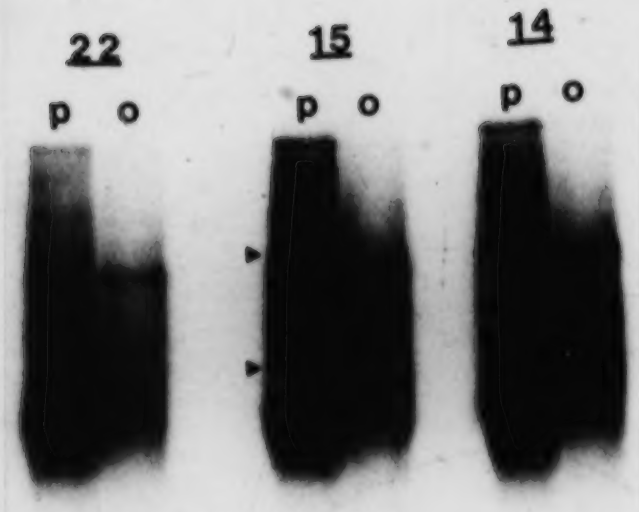


FIGURE 17. Clones 14, 15, and 22 detect a short PMA induced transcript in EL4.E1 cells. These three cDNA clones were used to probe three identical northern blots. All three probes gave a very similar pattern on northern analysis. Although there is a large amount of background, there appear to be three bands, which are best seen in 22. The lowest band is present only in lane p), and is completely absent in lane o).

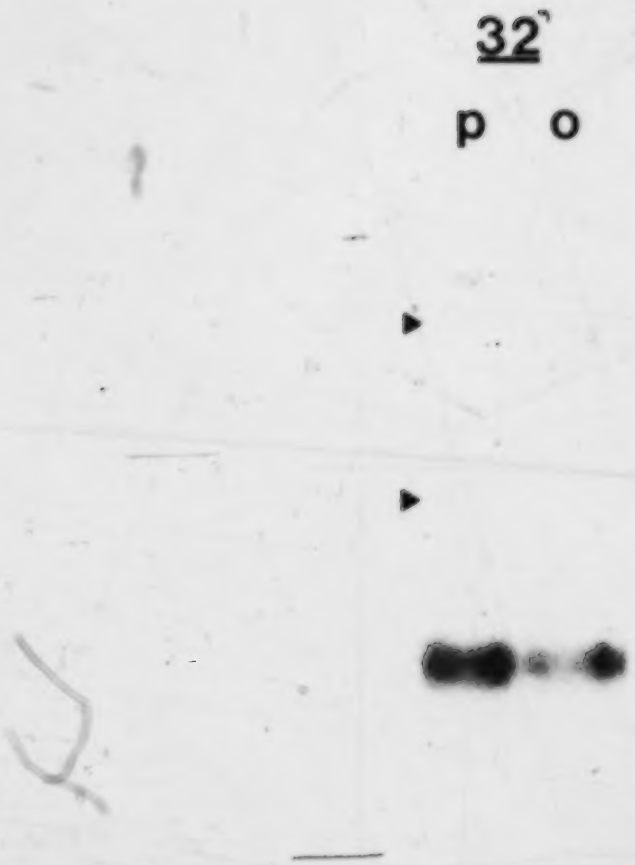


FIGURE 18. Clone 32 cDNA detects a transcript which is present in both PMA treated and untreated EL4.E1 cells. A single band of nearly equal intensity is seen in both lanes p) and o). The northern blot used in this experiment is identical to the one described in Figure 14.

shown). This demonstrates that the strategy using subtractive hybridization had also isolated known PMA-induced cDNAs. The characterisics of the other clones will be described in the following chapters.

CHAPTER VI

CHARACTERIZATION OF PMA-INDUCED cDNA CLONES FROM EL4

A. INTRODUCTION

The PMA-induced expression of mRNAs for IL2 and GM-CSF in EL4 cells is blocked if cells are stimulated in the presence of the immunosuppressive agent Cyclosporin A (CsA) (Chapter III). Furthermore, if cycloheximide (an inhibitor of protein synthesis) is added to EL4 cells several hours after PMA treatment, the accumulation of mRNAs for IL2 and GM-CSF is elevated significantly ('superinduction') (Paetkau *et al.*, 1985). This latter effect may in part be explained by the existence of a labile repressor protein which regulates the transcription of the induced mRNAs, and/or in part by the effect of cycloheximide (CHX) in stabilizing the induced mRNAs. The mechanism by which CsA blocks the induction of these transcripts is unknown at present. No matter what their mechanism of action, these two agents (CsA and CHX) appear to affect specifically the expression of only a small subset of all of the PMA-induced mRNAs in EL4 (Paetkau *et al.*, 1985), and they provide a means of classifying the various PMA-induced transcripts into groups according to their pattern of expression. For example, the two known lymphokine mRNAs (IL2 and GM-CSF) which are induced are both blocked in the presence of CsA. This has led to the hypothesis that perhaps all lymphokine mRNAs which are induced by PMA treatment of EL4 cells are blocked in the presence of CsA. The fact that the mRNAs for IL2 and GM-CSF are also 'superinduced' by CHX in EL4 has led to the suggestion that there is a correlation between superinducibility and suppression by CsA (Paetkau *et al.*, 1985).

It is likely that similar factors act to control the transcription and/or stability of those mRNAs which show a similar pattern of expression. Characterizing the

expression pattern of the various PMA-induced transcripts from EL4 may therefore be important in (eventually) discovering the mechanisms which control the concentration of these mRNAs within the cell. The first part of this chapter presents data on the effect of CsA and CHX on expression of the various PMA-inducible cDNA clones introduced in Chapter V.

To further characterize the unknown cDNA clones the primary nucleotide sequence of each clone was determined, and these data are presented and discussed in the second part of this chapter. In the absence of another strategy, this approach (simply sequencing the unknown clones) appeared to be the fastest way to discover more about the nature of the cDNA sequences and their corresponding mRNAs. Methods for determining DNA sequence (as described in Chapter II) are straightforward and relatively rapid, and yield exact data which can be used to search a computer data-base to identify sequences (DNA or protein) which have already been reported. If a clone does not correspond to a known sequence, its (DNA) sequence is nevertheless required to predict the sequence of the corresponding protein.

B. PATTERNS OF EXPRESSION OF PMA-INDUCED cDNA CLONES

Figure 19 shows the autoradiograph which resulted when a northern blot of total cellular RNA from EL4 cells was probed with labelled insert from clone 12 (this clone was shown to be IL2 as described in chapter V). As described above, the IL2 transcript is absent in control cells (lane o]), induced in PMA-treated cells (lane p]), and blocked in the presence of PMA and CsA (lane c]). Furthermore, by adding CHX ten hours after PMA treatment (lane x]), the transcript appears to be 'superinduced' about 4-fold (compare lanes x] and p]). This blot serves as a positive control for a number of identical blots which follow.

To ensure that equivalent amounts of intact mRNA from each lane were

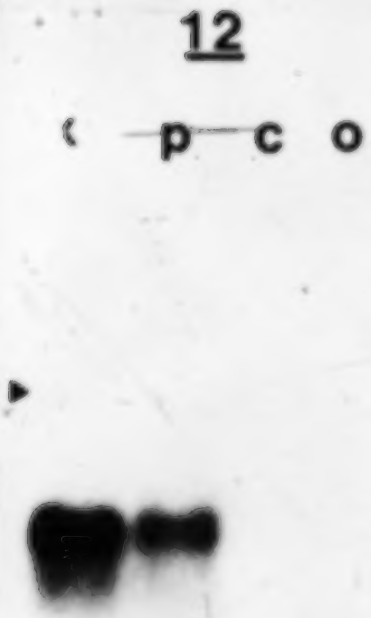


FIGURE 19. Expression pattern of the transcript detected by clone 12. Total cellular RNA was isolated from EL4.E1 cells, separated by electrophoresis on a 1% agarose-formaldehyde gel, and transferred to nitrocellulose. This northern blot was probed with labelled cDNA insert from clone 12, using standard methods. Twenty µg of total RNA was loaded in each lane. The RNA was from EL4.E1 cells treated as follows: lane x), PMA (20 ng/ml) for 16 hours, with cycloheximide (20 µg/ml) for the last 6 hours of culture; lane p), PMA (20 ng/ml) for 16 hours; lane c), PMA (20 ng/ml) and CsA (50 ng/ml) for 16 hours; lane o), untreated EL4 cells. The arrows mark the position of the 18S and 28S mitochodrial RNA bands.

transferred to the northern blots used in this study a second control was also done. For this, a northern blot (identical to that used in the Figure 19) was probed with labelled cDNA insert from clone 32 (Figure 20). As described in chapter V, the mRNA is present in equivalent amounts in both control cells (lane o]) and EL4 cells treated with PMA (lane p]). Furthermore, neither CsA nor CHX affects its expression, since bands of equal intensity appear in all four lanes of Figure 20.

Figure 21 shows the results of a northern blot which has been probed with clone 25 cDNA insert. The transcript detected is clearly induced by PMA (a band is absent in lane o], but present in lane p]), but levels of expression are unaffected by either CsA or CHX (the bands in lanes x], p], and c] are of equivalent intensity). Figure 22 shows a similar result for clone 154, but again demonstrates that a small amount of this transcript is also made in untreated EL4 cells (a very faint band can be detected in lane o]; compare Figure 15, Chapter V).

Figure 23 shows the results of a northern blot which has been probed with clone 14. A series of longer transcripts produces a series of indistinct bands (the northern blot which was made with total RNA, rather than poly A+ RNA, and appears to give poorer resolution of the various transcripts). In general these longer transcripts are slightly induced in PMA-treated cells. The outstanding feature of these blots, however, is a very dark band which is present in lane p) but absent in lane o). This short PMA-induced transcript is blocked in the presence of CsA (the short, dark band is absent in lane c]), and appears to be relatively unaffected by CHX (the bands are of equal intensity in lanes p] and x]). The nature and sequence of this transcript will be described in much more detail in Chapter VIII.

The final northern blot in this series (Figure 24) shows the results when clone 55 cDNA insert was used as a probe. The transcript detected is induced by PMA (a band is absent in lane o] and present in lane p]). The surprising result is that this transcript

FIGURE 20. Expression pattern of the transcript detected by clone 32. The northern blot used in this experiment is identical to the one described in Figure 19.

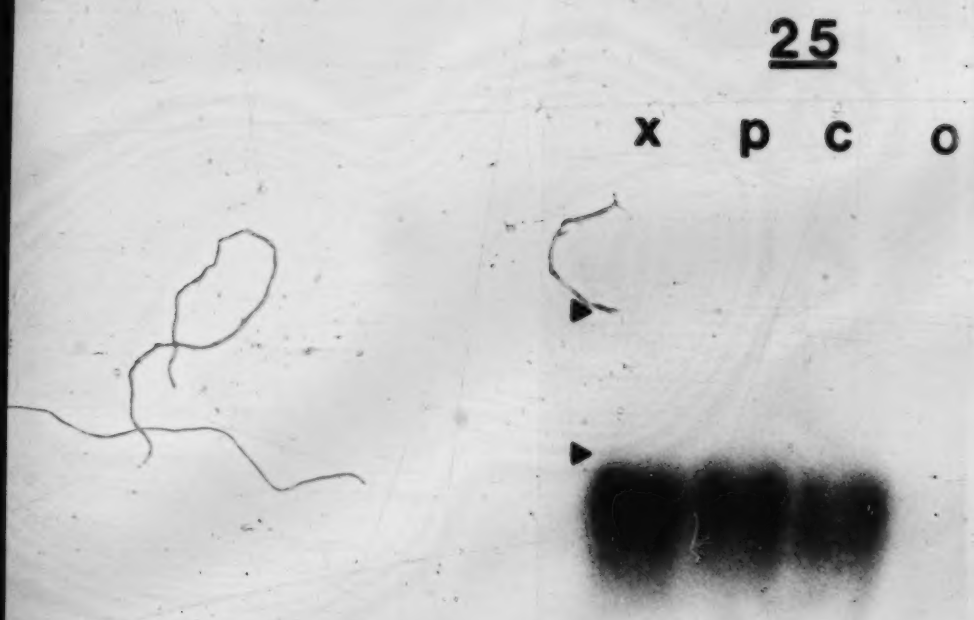


FIGURE 21. Expression pattern of the transcript detected by clone 25. The northern blot used in this experiment is identical to the one described in Figure 19.

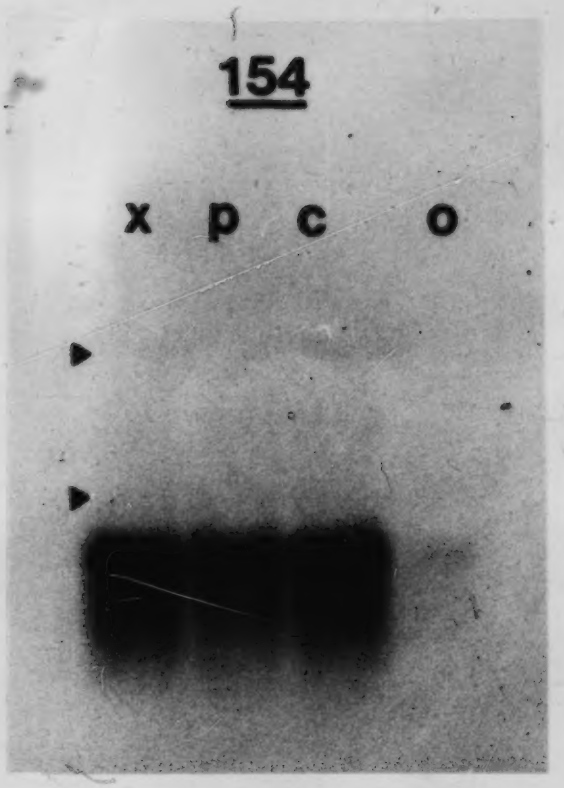


FIGURE 22. Expression pattern of the transcript detected by clone 154. The northern blot used in this experiment is identical to the one described in Figure 19.

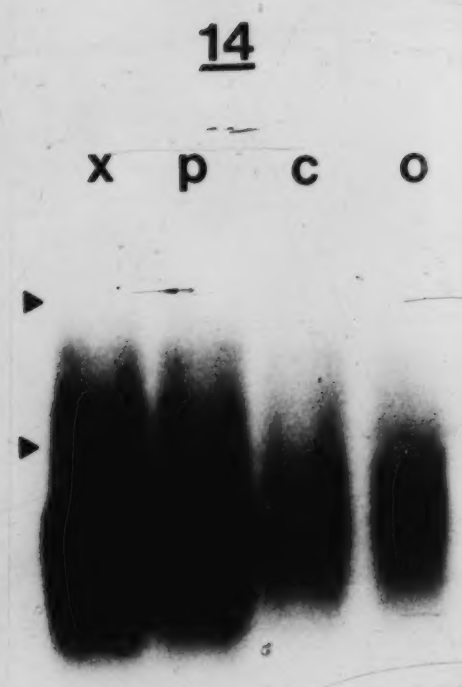


FIGURE 23. Expression pattern of the transcript detected by clone 14. The northern blot used in this experiment is identical to the one described in Figure 19.

FIGURE 24. Expression pattern of the transcript detected by clone 55. The northern blot used in this experiment is identical to the one described in Figure 19.

is not blocked by CsA, but rather induced to levels much higher than those seen with PMA alone (the band appears much darker in lane c] than in lane p]). Levels of the mRNA appear to be unaffected by CHX (the band in lanes p] and x] are of equal intensity). The results for this clone are confirmed and extended in Figure 25. This figure shows the results of a 'cytodot' (White and Bancroft, 1982), which was made by blotting total cytoplasmic RNA (from EL4 cells) directly onto nitrocellulose and then probing with clone 55. Again the transcript detected is induced in PMA-treated cells (a spot is detected in column p], but not in column o]). The transcript is induced to even higher levels in the presence of PMA and CsA (column p/c] shows a much larger and darker spot in the first row). In fact the transcript appears to be induced about 16-fold in the presence of PMA and CsA, as compared to PMA alone, since 4 two-fold dilutions can be made in the former RNA to give an equivalent signal in both samples (*i.e.* the fifth spot in column p/c] gives about the same signal as the first spot in column p]). The presence of CsA alone has virtually no effect on the induction of the transcript, since column c) gives nearly the identical signal to that seen in column o).

The expression patterns of the various cDNA clones tested are summarized in Table 7. The expression pattern of clone 11 (Chapter V) was not tested, since initial northern analysis failed to detect a single predominant PMA-induced transcript (Figure 16). The results of the various northern blots described above serve to confirm that the cDNA clones isolated represent PMA-induced transcripts present in EL4 cells. The original northerns used in Chapter V were made from the same poly A^+ RNA used to generate the subtracted library and subtracted probe, and thus they represent an internal control for the cloning and subtraction process. The northern blots described in this chapter were made from a completely different preparation of total RNA, and thus they provide an 'external control' for the same process.

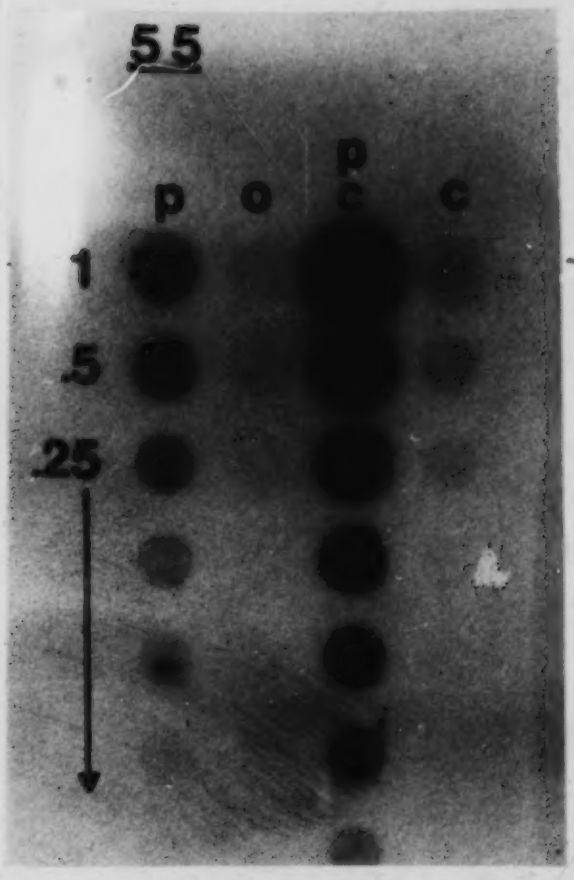


FIGURE 25. Expression pattern of the transcript detected by clone 55, using the cytodot method. Total cytoplasmic RNA was isolated from EL4.E1 cells, and fixed to nitrocellulose using a 96 hole Minifold apparatus (Schleicher and Schuell). For each RNA sample, serial two-fold dilutions were made and applied to the nitrocellulose, so that RNA from 1 million cells is fixed to the first (4mm) spot, RNA from 0.5 million cells is applied to the second spot, and so on. The various samples of RNA are from EL4.E1 cells treated as follows: column p), PMA (20 ng/ml) for 16 hours; column o), untreated cells; column p/c), PMA (20 ng/ml) and CsA (50 ng/ml) for 16 hours; column c), CsA (50 ng/ml) alone for 16 hours.

C. PRIMARY NUCLEOTIDE SEQUENCES OF THE PMA-INDUCED CLONES

The restriction map and a partial sequence of clone 12 was shown to be identical to that of the published sequence of murine IL2 (Fuse *et al.*, 1984), and it will not be discussed further. The sequence of clone 14 demonstrated that the transcript was derived from the *env* and LTR region of mouse mammary tumor virus (MMTV), and these results are presented in Chapter VIII.

Clone 32 did not appear to be induced by PMA, nor was it affected by CsA or CHX. We would not expect to isolate a clone with this expression pattern using the subtractive strategy outlined in chapter V, and it is likely that its isolation is an artifact or false-positive of the subtractive cloning and selection process. What is slightly surprising is the fact that the sequence appeared twice in the original 136 clones selected (Table 7), suggesting that there is some systematic (but completely unknown) reason why it was consistently isolated during the subtractive process. The clone 32 cDNA was used as a control in northern blots (as demonstrated above), and for this reason the primary sequence was determined (Figure 26). The sequence shows a number of interesting features. The entire sequence is an open reading frame, and codes for a protein which is nearly identical to a portion of the human pro-ubiquitin protein described by Lund *et al.* (1985). The mouse sequence differs from the human by 25 single base changes (bases which differ are indicated on Figure 26), but the corresponding protein sequence is identical to the human sequence except for one conservative change (ala to gly at position 57). The human sequence has an *Eco* RI site about 100 bases from the 5' end, and this site is probably also present in the mouse sequence, and suggests that the 5' end of the sequence was lost during the subcloning process. The total length of the original cDNA insert is estimated to be ≈500 bp (only a partial sequence is given in Figure 26), and this is consistent with the fact that the cDNA insert detects a single short transcript of about the same size. Although the

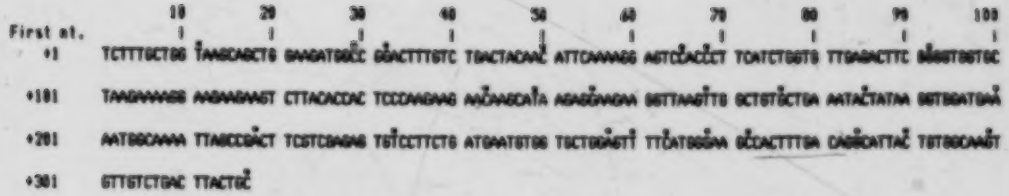


FIGURE 26. Partial nucleotide sequence of clone 32 cDNA. This sequence can be aligned with the published cDNA sequence of human pro-ubiquitin (nucleotides 118 to 435 [Lund *et al.*, 1985]). Bases which are different between the two sequences are indicated by a 'dot' on the clone 32 sequence. This DNA fragment was sequenced on both strands, using a single M13 template for each direction; thus the sequencing strategy was as follows: upper strand 1-317; lower strand 317-1, where the upper strand refers to the one shown in the figure. (This abbreviated notation will also be used to describe the sequencing strategy used to determine the DNA sequences shown in subsequent figures.)

cDNA inserts were size selected for DNAs ≥ 800 bp when the original library was constructed, this process of size selection is apparently not perfect.

Figure 27 shows the sequence of a portion of clone 25. This sequence is almost identical (bases 1 to 712) to that of mouse carbonic anhydrase II described by Curtis *et al.* (1983), and contains coding information for the 112 C-terminal amino acids in this enzyme (the complete enzyme is about 260 amino acids long). The sequence diverges from that of carbonic anhydrase at base 713, and the string of C's at this position indicate that another small fragment of (completely unrelated) DNA was probably ligated to the carbonic anhydrase cDNA during the ligation with *Eco* RI linkers. This cloning artifact has been described by other investigators (C. Benoist, unpublished observations). In any case, the major part of clone 25 contains a cDNA sequence capable of detecting a mouse carbonic anhydrase II transcript, and such a transcript is induced by PMA in EL4 cells, but unaffected by CsA or CHX (Figure 21). Clone 25 contains only part of the cDNA sequence for carbonic anhydrase II; the full-length clone reported by Curtis *et al.* (1983) is 1480 bp in length. The size of the transcript detected by clone 25 in EL4 has not been determined exactly, but it appears to be reasonably close to the 1480 base transcript predicted by the full-length cDNA (Figure 21). Carbonic anhydrase II is the second most abundant protein in erythrocytes, and it has been shown that this transcript is increased 2-fold (from 0.04% to 0.077% of the poly A^+ RNA) when mouse erythroleukemia cells are induced to differentiate by treatment with dimethyl sulfoxide (Curtis, 1983). Whether or not the carbonic anhydrase transcript is present in lymphoid cells has apparently not been determined by these investigators.

The complete nucleotide sequence of clone 154 is given in Figure 28. The sequence has a single large open reading frame which begins at base 62 and ends at base 520. Figure 29 shows the predicted amino-acid sequence of the corresponding

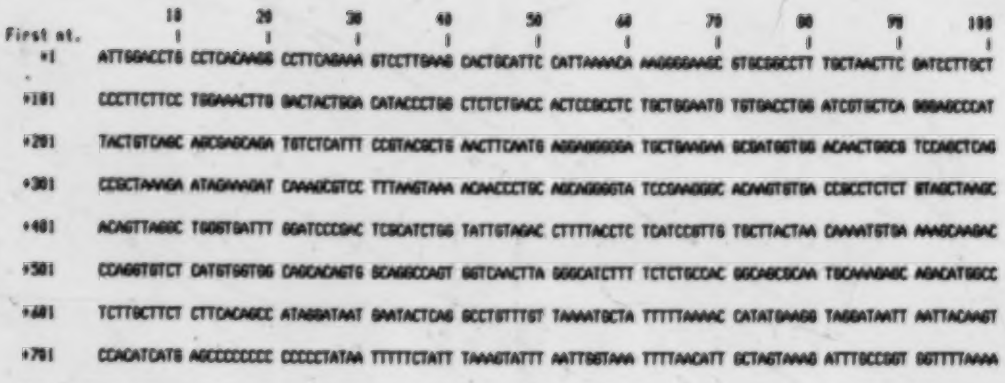

First at.
10 20 30 40 50 60 70 80 90 100
+1 ATTGGACCTG CCTCACAAGG CCTTCAGAAA GTCCTTGAAG CACTGCATTC CATTAAAACA AAGGGGAAGC GTGCGGCCTT TGCTAACTTC GATCCTTGCT
+101 CCCTTCTTCC TGGAAACTTG GACTACTGGA CATACCCTGG CTCTCTGACC ACTCCGCCTC TGCTGCAATG TGTGACCTGG ATCGTGCTCA GGGAGCCCAT
+201 TACTGTCAGC AGCGAGCAGA TGTCTCATTT CCGTACGCTG AACTTCAATG AGGAGGGGGA TGCTGAAGAA GCGATGGTGG ACAACTGGCG TCCAGCTCAG
+301 CCGCTAAAGA ATAGAAAGAT CAAAGCGTCC TTTAAGTAAA ACAACCCTGC AGCAGGGGTA TCCGAAGGGC ACAAGTGTGA CCGCCTCTCT GTAGCTAAGC
+401 ACAGTTAGGC TGGGTGATTT GGATCCCGAC TCGCATCTGG TATTGTAGAC CTTTTACCTC TCATCCGTTG TGCTTACTAA CAAAATGTGA AAAGCAAGAC
+501 CCAGGTGTCT CATGTGGTGG CAGCACAGTG GCAGGCCAGT GGTCAACTTA GGGCATCTTT TCTCTGCCAC GGCAGCGCAA TGCAAAGAGC AGACATGGCC
+601 TCTTGCTTCT CTTCACAGCC ATAGGATAAT GAATACTCAG GCCTGTTTGT TAAAATGCTA TTTTTAAAAC CATATGAAGG TAGGATAATT AATTACAAGT
+701 CCACATCATG AGCCCCCCCC CCCCCTATAA TTTTTCTATT TAAAGTATTT AATTGGTAAA TTTTAACATT GCTAGTAAAG ATTTGCCGGT GGTTTTAAAA
+801 AA

FIGURE 27. Partial nucleotide sequence of clone 25 cDNA insert. This sequence can be aligned with the published sequence for mouse carbonic anhydrase II (beginning at the published nucleotide 472 [Curtis *et al.*, 1983]). The DNA contains two useful restriction sites; *Pst* I at 345, and *Bam* HI at 415. Sequencing strategy: upper, 1-330,348-706; lower 802-409, 415-129, 345-20. (*i.e.* This DNA fragement has not been completely sequenced over the entire length of each strand.)

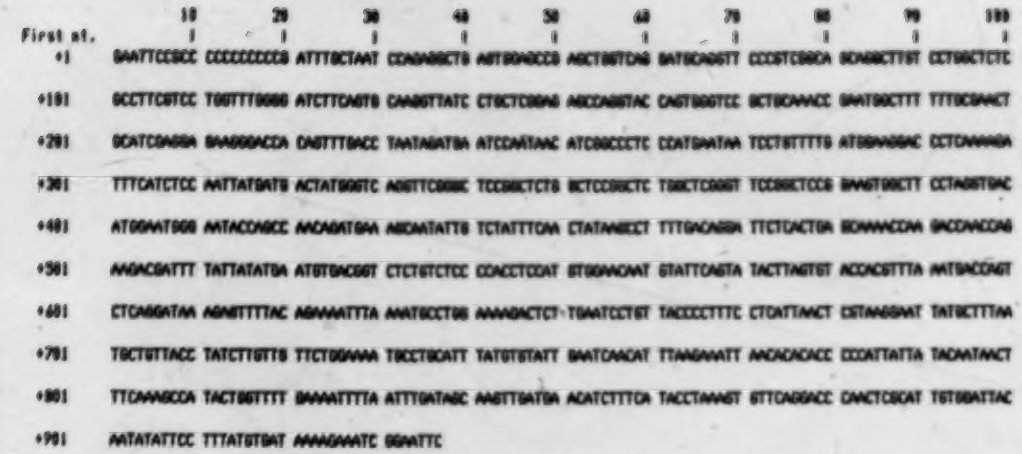


FIGURE 28. Complete nucleotide sequence of clone 154 cDNA insert. Important restriction sites include *Acc* I at 568, *Ava* I at 364, and *Rsa* I at 157 and 579. Sequencing strategy: upper, 1-370, 7-198, 223-488, 364-710, 367-548, 509-731, 568-937, 579-872; lower, 937-551, 780-510, 720-522, 713-440, 681-458, 568-155, 479-334, 364-43, 277-6, 157-1.

met-gln-val-pro-val-gly-ser-arg-leu-val-leu-ala-leu-ala-phe-val-leu-val-trp-gly-
ser-ser-val-gln-gly-tyr-pro-ala-arg-arg-ala-arg-tyr-gln-trp-val-arg-cys-lys-pro-
asn-gly-phe-phe-ala-asn-cys-ile-glu-glu-lys-gly-pro-gln-phe-asp-leu-ile-asp-glu-
ser-asn-asn-ile-gly-pro-pro-met-asn-asn-pro-val-leu-met-glu-gly-pro-ser-lys-asp-
phe-ile-ser-asn-tyr-asp-asp-tyr-[gly-ser-gly-ser-gly-ser-gly-ser-gly-ser-gly-ser-
gly-ser-gly-ser-gly-ser-gly-ser-gly-ser-gly]phe-leu-gly-asp-met-glu-trp-glu-tyr-gln-pro-
thr-asp-glu-ser-asn-ile-val-tyr-phe-asn-tyr-lys-pro-phe-asp-arg-ile-leu-thr-glu-
gln-asn-gln-asp-gln-pro-glu-asp-asp-phe-ile-ile-STOP

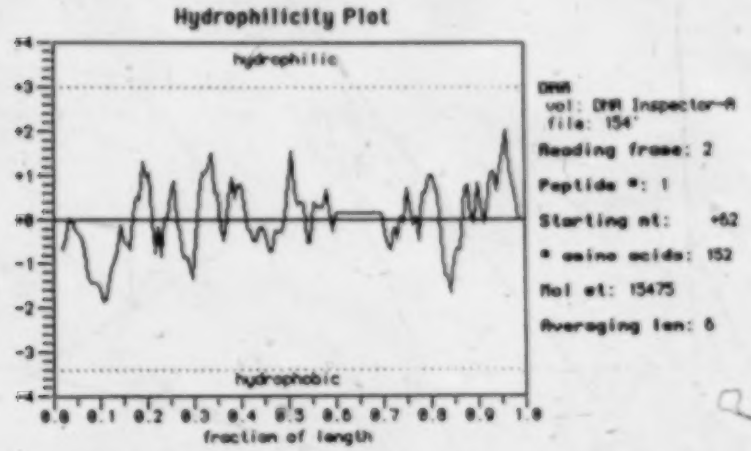


FIGURE 29. Amino acid sequence and Hopp and Woods hydrophilicity analysis of the protein coded by clone 154. The protein has an unusual stretch of 21 residues which alternate gly-ser-gly-ser, and these are enclosed in brackets. Residues which differ between this sequence and the the published sequence for rat chondroitin sulfate core protein are underlined (Bourdon *et al.*, 1986). The hydrophilicity plot demonstrates that the protein has an N-terminal hydrophobic leader sequence (15 amino acids in length).

protein. The protein has an unusual stretch of 21 amino acids with alternating glycine and serine residues (enclosed in brackets). The deduced protein sequence for rat chondroitin sulfate core protein (from the cDNA sequence described by Bourdon *et al.*, 1986) has a similar central 49 amino acid region composed of alternating serine and glycine residues. Furthermore these two proteins are highly homologous over most of their length (residues which are not conserved between the mouse and rat proteins are underlined in Figure 29). The two cDNA sequences are also nearly identical over large stretches of the N-terminal, C-terminal, and 3' untranslated regions (≈85% conservation at the nucleotide level). These observations suggest that the clone 154 sequence codes for a proteoglycan molecule which is present in mouse lymphocytes. Proteoglycans consist of a core protein with multiple glycosaminoglycan chains attached via *O* -glycosyl linkage to serine residues. These molecules are present at the cell surface, interact with other extracellular matrix proteins, and are thought to be involved in cellular adhesion (Bourdon *et al.*, 1985). EL4 cells normally grow in grape-like clusters which are non-adherent. However, following PMA treatment the cells become dissociated, take on a more polygonal shape, and become firmly adherent to the bottom of the culture flask. Increased synthesis of proteoglycan during PMA treatment of EL4 may in part mediate the change in morphology seen in these cells following PMA-stimulation.

Figure 30 gives the primary nucleotide sequence of clone 55. The sequence shows a single large open reading frame of 133 amino acids which begins at base 69 and ends at base 470. The corresponding protein sequence is shown in Figure 31. The DNA and derived protein sequences have not been previously reported. The size of the cDNA sequence (1510 bp) corresponds reasonably well to the size of the transcript detected on northern blots (Figures 14 and 24). The cDNA sequence reported here is likely almost full length, since when the cDNA insert is used to transcribe RNA

FIGURE 30. Complete nucleotide sequence of clone 55 cDNA insert. An eight bp sequence which is found in the 3' untranslated region of a number of lymphokine cDNA's is also found three times in this sequence (underlined; see text). Important restriction sites include *Sst* I at 187, *Nde* I at 566, *Rsa* I at 310 and 445, *Pvu* II at 107, 171, and 1229. Sequencing strategy: upper, 1-211, 156-359, 187-461, 315-546, 434-643, 445-781, 566-919, 666-855, 750-947, 876-1082, 968-1189, 1055-1288, 1229-1510, 1238-1510; lower, 1510-1104, 1377-1061, 1146-946, 996-796, 878-678, 800-600, 621-472, 566-305, 331-41, 310-31, 187-1, 88-1.

met-arg-arg-met-leu-leu-his-leu-ser-val-leu-thr-leu-ser-cys-val-trp-ala-thr-ala-
met-glu-ile-pro-met-ser-thr-val-val-lys-glu-thr-leu-thr-gln-leu-ser-ala-his-arg-
ala-leu-leu-thr-ser-asn-glu-thr-met-arg-leu-pro-val-pro-thr-his-lys-asn-his-gln-
leu-cys-ile-gly-glu-ile-phe-gln-gly-leu-asp-ile-leu-lys-asn-gln-thr-val-his-gly-
gly-thr-val-glu-met-leu-phe-gln-asn-leu-ser-leu-ile-lys-lys-tyr-ile-asp-arg-gln-
lys-glu-lys-cys-gly-glu-glu-arg-arg-arg-thr-arg-gln-phe-leu-asp-tyr-leu-gln-glu-
phe-leu-gly-val-met-ser-thr-glu-trp-ala-met-glu-gly-STOP

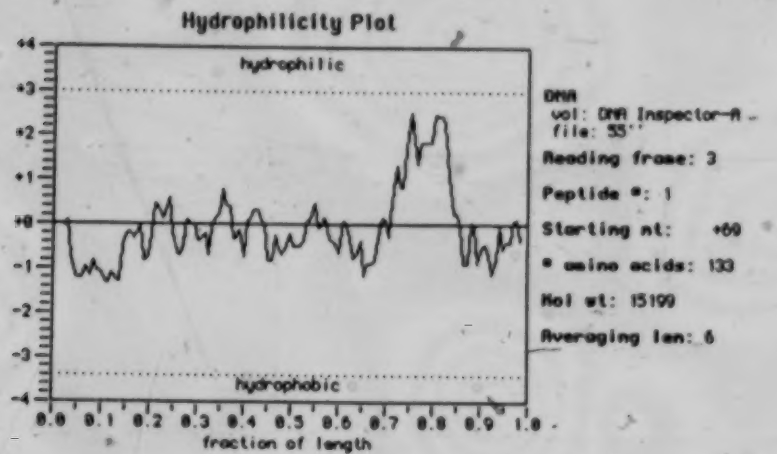


FIGURE 31. Amino acid sequence and Hopp and Woods hydrophilicity analysis of the protein coded by clone 55. The protein has 3 potential N-linked glycosylation sites, and these are underlined. The hydrophilicity analysis demonstrates that the protein has a N-terminal hydrophobic signal sequence, and the proteolytic cleavage which yields the mature protein probably occurs between ala-18 and thr-19.

(using an upstream T7 promoter and T7 polymerase) and the RNA run on a denaturing gel, it runs in almost the same position as does the transcript detected in a northern blot of poly A^+ RNA from EL4 cells (D. Talbot, personal communication). The nucleotide sequence of clone 55 has a very large 3' untranslated region, a feature which has been noted for a number of lymphokine mRNAs. Caput *et al.* (1986) have described a consensus sequence (TTATTTAT) which is found in the 3' untranslated region of mRNAs coding for a number of interleukins, interferons, and CSFs; it appears to be specific to the mRNAs coding for immunoregulatory proteins. This conserved octamer is also found in the 3' untranslated region of the clone 55 sequence. It occurs 3 times (underlined); at base 584 (one mismatch), base 1458 (one mismatch) and base 1484 (perfect match). The deduced protein sequence for clone 55 shows a probable signal sequence of met-arg-arg followed by 18 hydrophobic amino acids and then glu (see hydrophilicity plot, Figure 31). Such a signal sequence makes it highly probable that the 55 protein is secreted by the cell (J. Jenson, personal communication). The protein also contains three potential N-glycosylation sites (with the sequence asn-x-thr/ser) at asparagines 46, 75, and 89. This cDNA sequence has a number of features that makes it reasonable to suggest that the transcript may code for a lymphokine or other immunoregulatory molecule. Additional data concerning clone 55 is presented in chapter VII.

D. DISCUSSION

The characteristics of the various PMA-induced cDNA clones introduced in Chapter V have been outlined in this chapter. These data are summarized in Table 7, along with the relative abundance of the various sequences in the original 136 clones isolated. In terms of expression patterns, two clones (25 and 154) appear to be unaffected by CsA or CHX, two are blocked in the presence of CsA (clones 12 and

TABLE 7

SUMMARY OF PMA-INDUCED cDNA CLONES FOUND IN EL4

Prototype clone	Number found in 136 clones isolated	Expression Pattern	Identity of sequence
14	99	induced by PMA blocked by CsA unaffected by CHX	short MMTV transcript
11	18	large number of bands on northern analysis, all inducd by PMA; nfc	unknown
55	5	induced by PMA superinduced by CsA unaffected by CHX	unknown (possible new lymphokine?)
12	4	induced by PMA blocked by CsA superinduced by CHX	murine IL2
154	2	induced by PMA unaffected by CsA unaffected by CHX	core protein for chondroitin sulfate
25	1	induced by PMA unaffected by CsA unaffected by CHX	murine carbonic anhydrase II
32	2	unaffected by PMA, CsA, or CHX; why isolated???	murine pro-ubiquitin
Others	5	no subclonable *Eco* RI inserts, or no detectable band on northern analysis	likely cloning artifacts

nfc: not further characterized
PMA: phorbol-12-myristate-13-acetate
CsA: cyclosporin A
CHX: cycloheximide

14), and one is 'superinduced' in the presence of CsA (clone 55). Clone 12 was 'superinduced' by CHX, whereas clones 14 and 55 were unaffected by CHX. Clone 11 was not analyzed since it did not cross-react with a single obvious PMA-induced transcript on northern analysis. Primary sequence data have allowed five of the cloned sequences to be identified (clones 14, 25, 12, 154 and 32), and one of the sequences (clone 55) cannot be found in the literature. One of the clones isolated is a known lymphokine (clone 12 is IL2), and one (clone 55) is potentially a new lymphokine. The only known PMA-induced mRNA from EL4 cells which was not found in this study was GM-CSF. This transcript is at least five-fold less abundant than the IL2 transcript (Chapter IV). Since IL2 occurred only 4 times in the original 136 clones isolated, it is reasonable to suggest that if a larger number of clones had been isolated, GM-CSF may well have been among the clones selected.

Almost three-quarters of the clones isolated belonged to the 'clone 14' family and were derived from the retrovirus MMTV. These results are in reasonable agreement with those of Kwon and Weissman (1984), who found that half of their PMA-induced clones from EL4 were derived from MMTV. These observations demonstrate one of the serious problems with using a tumor cell line such as EL4 to clone genes which are expressed in normal T-lymphocytes. Tumor cells may express genes which are peculiar to their state, and the corresponding transcripts may constitute a large proportion of all the mRNAs of interest, making the isolation of more 'normal' cDNAs considerably more difficult. On the other hand, the subtractive strategy requires a relatively large amount of RNA from cells of clonal origin, and these amounts of RNA can be readily obtained only from cells (such as tumor cell lines) which can be easily cloned and grow rapidly in culture. The clone 14 retroviral transcript may provide a useful model to study the regulation of other PMA-inducible genes such as lymphokines (see Chapter VIII), so that isolation of transcripts peculiar to tumor cells is

not necessarily an undesirable result providing that these mRNAs have an interesting expression pattern. In any case, if additional PMA-induced cDNA sequences from EL4 were to be sought by using the subtractive strategy outlined in Chapter V, a very important step would be the elimination of the MMTV sequences (and other highly abundant induced sequences such as clone 11) from the subtracted cDNA probe.

CHAPTER VII

FURTHER CHARACTERIZATION OF THE 'CLONE 55' TRANSCRIPT

A. INTRODUCTION

The nucleotide sequences of the various PMA-induced cDNA clones made it possible to identify all but one of the clones; the clone 55 sequence did not correspond to any known mRNA or protein (Chapter VI). The sequence of this clone has several features which suggest that the corresponding transcript may code for a lymphokine or other immunoregulatory molecule. If the putative lymphokine plays some physiological role in the immune system, we would expect that the protein would be made by at least some normal T-helper cells, and that the corresponding transcript could be found in such cells. To test this idea we initially screened a number of other mouse cell lines of immune origin for the presence of the clone 55 transcript. In this chapter we show that the 55 transcript is found in an antigen-dependent T-helper cell line, and the mRNA is induced in these cells when they are stimulated with mitogen.

To further investigate the function of the clone 55 product it would be useful to obtain the protein coded for by this transcript. A number of eucaryotic expression systems have been developed which enable the investigator to produce the protein coded for by any given cDNA molecule, provided the cDNA is complete (*i.e.* contains the entire protein coding sequence, including the start codon). To obtain the protein coded for by the clone 55 cDNA sequence, the cDNA was placed in an SV40-derived expression vector and transfected into COS monkey fibroblasts (Chapter II). The resulting material was tested for growth promoting or lymphokine activity in a number of *in vitro* bioassays, and these results are presented in the second part of this chapter. As a positive control the same expression system was used to produce mouse IL2,

mouse GM-CSF, and human IL2, using the cDNA molecules cloned previously (Chapter IV).

B. RESULTS

1. *Expression Pattern of the Clone 55 Transcript in Other Lymphoid Cell Lines*

Table 8 summarizes the results which were obtained when a number of cell lines were screened for the presence of the clone 55 transcript (using the cytodot method; Chapter VI). For each case the cells were assayed either in the resting state, or following mitogen stimulation. Only two cell lines, EL4 and CH2.4 gave detectable levels of the 55 transcript, and in each case the transcript was only present in cells treated with mitogen (PMA and/or Con A). For comparison the same cells were also probed for the presence of the clone 154 transcript (chondroitin sulfate). This transcript was constitutively expressed in almost all the cell lines, and was induced by mitogen in only a few cases.

The upper part of Figure 32 shows a northern blot which confirms and extends the results for the CH2.4 cells. The clone 55 transcript is absent in resting cells (lane O]), and present in mitogen-treated cells (lane A]). A surprising result is that the 55 transcript is reduced markedly when CH2.4 cells are treated with mitogen in the presence of CsA (lane C]); this is the opposite result from what was found in EL4 cells (Chapter VI). The size of the 55 transcript in CH2.4 cells (lane A]) is identical to that seen in PMA-treated EL4 cells (lane E]). The cytodot analysis revealed that 'round cells' do not make the 55 transcript (Table 8), and this result is confirmed in lane R). The lower panel in Figure 32 shows the same blot probed with the clone 154 insert. In contrast to EL4, the 154 transcript is expressed at high levels in resting CH2.4 cells, and is only slightly induced following mitogen stimulation. The bands are of equal intensity in lanes A) and C), and since CsA does not affect the 154 transcript, this implies that equal quantities of RNA were present in both lanes of the blot. This latter

TABLE 8

SUMMARY OF CELL LINES SCREENED FOR EXPRESSION OF THE 'CLONE 55 TRANSCRIPT' USING CYTODOTS

Name of cell line	Type of cell	Transcripts detected and expression pattern		
		murine IL2	clone 55	clone 154
EL4.E1	T-helper	inducible*	inducible	inducible
CH2.4*	T-helper	absent	inducible	high con.
P815	mastocytoma	absent	absent	low con.
BW5147	T-cell (?helper)	absent	absent	inducible
P388D1	macrophage	absent	absent	low con.
L5178Y	T-cell	absent	absent	low con.
L1210	B-cell	absent	absent	low con.
SL25	T-cell	absent	absent	high con.
MTL2.8.2*	cytotoxic T-cell	absent	absent	absent (?)
'Round cells'*	unknown	absent	absent	low con.
B1.1#	T-helper hybridoma	absent	absent	high con.
Con A treated splenocytes	mixture of cell types	absent(?)[9]	absent	n.d.

*all cells were tested both in the resting state and after mitogen stimulation
high con.: high constiutitive expression; low con.: low constituitive expression
*cell lines developed by C. Havele, University of Alberta
#an anti poly-18 cell line described by Fotedar *et al.* (1985);
The remaining cell lines are described in the American Type Culture Collection catalogue;
[9]IL2 transcripts were apparently below detectable levels, although IL2 is produced in Con A treated splenocytes.
n.d.: not determined

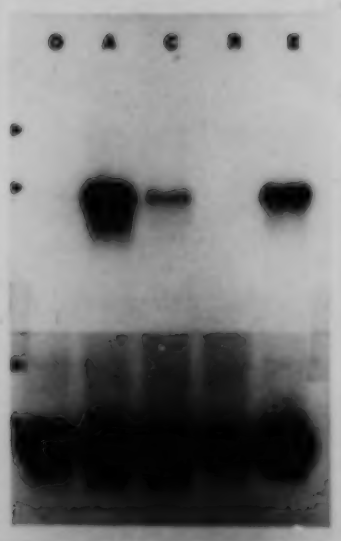


FIGURE 32. Expression pattern of the transcript detected by clone 55 in the helper T cell line CH2.4. Total cellular RNA was isolated from cells, separated by electrophoresis on a 1% agarose-formaldehyde gel, and transferred to nitrocellulose. The resulting northern blot was probed with labelled cDNA insert from clone 55 (top part of the figure), washed, and re-probed with labelled cDNA insert from clone 154 (bottom part of the figure). Eight μg of total cellular RNA was loaded per track. The various RNA samples are from cells treated as follows: lane O), resting CH2.4 cells; lane A), CH2.4 cells cultured with Con A (2 μg/ml) for 16 hours; lane C), CH2.4 cells cultured with Con A (2 μg/ml) and CsA (300 ng/ml) for 16 hours; lane R), 'round cells' cultured with Con A (2 μg/ml) for 16 hours; lane E), EL4.E1 cells cultured with PMA (20 ng/ml) for 16 hours. The arrows indicate the location of the 18S and 28S ribosomal RNA bands.

observation rules out any trivial explanation for the difference between lanes A) and C) seen in the upper part of the figure.

2. *Expression of the Clone 55 Protein in COS Cells and Bioassay for Function*

The clone 55 cDNA insert is very likely full length for the reason outlined in Chapter VI. Furthermore, the apparent coding region indicates that the clone 55 protein is secreted, since the N-terminus codes for a signal sequence. This latter observation makes it even more likely that the clone 55 cDNA is full length, and suggests that if the protein were expressed in COS cells it would be secreted into the culture media. The clone 55 cDNA insert was ligated into an expression vector containing the SV40 promoter-enhancer region, and these constructs were transfected into COS cells using DEAE-dextran followed by a glycerol shock (the vector and transfection method are described in Chapter II). The cDNA insert was ligated in both the forward and reverse orientations (relative to the promoter-enhancer element), and both constructs were transfected in parallel. Supernatants resulting from the 'reverse oriented' construct served as negative controls for the COS cell supernatants. In such an expression system it is difficult to design a perfect positive control to show that the COS cells transfected with the 'forward oriented' construct are actually secreting the clone 55 protein, since there is no known bioassay for the protein, and the quantities of protein produced make biosynthetic labelling experiments very difficult. For our purposes, the positive control for these experiments involved making similar constructs using the mouse IL2, mouse GM-CSF, and human IL2 inserts (in both forward and reverse orientations), transfecting these constructs in parallel with the clone 55 constructs, and showing that the COS cell supernatants from these known lymphokine cDNAs gave a positive result in the appropriate bioassay. The results of these experiments are shown in Table 9. Both the mouse and human IL2 constructs produced significant amounts of

TABLE 9

RESULTS OF BIOASSAY FOR LYMPHOKINE ACTIVITY IN TRANSFECTED COS CELL SUPERNATANTS

cDNA insert used for transfection	pcEXV-3 subclone no.	insert orientation	IL2 activity (units/ml)	GM-CSF* activity (U/ml)
mouse IL2(clone 12)	18	forward	383	<bgd.
	19	reverse	<bgd.	<bgd.
mouse GM-CSF	10	forward	<bgd.	512
	9	reverse	<bgd.	<4
human IL2	1	forward	212	<bgd.
	2	reverse	<bgd.	<bgd.
clone 55	102	forward	<bgd.	<bgd.
	103	reverse	<bgd.	<bgd.

The IL2 bioassay was done as described in Chapter II.
*GM-CSF bioassay was performed by L. Gilbert, University of Alberta, Edmonton, as described (Branch and Gilbert, in press); 1 unit of GM-CSF gives 50% of the maximal response.
<bgd.: less than or equal to background levels of activity in the standard bioassay.

IL2 when the COS cell supernatants were bioassayed for IL2. Similarly, the mouse GM-CSF construct produced the expected lymphokine. The COS cell supernatants from the clone 55 constructs gave a negative result in both the IL2 and GM-CSF bioassays. The 55 supernatants were tested in a large number of other bioassays; the results of these investigations are summarized in Table 10. In short, no lymphokine-like or growth promoting activity has been discovered in these supernatants.

C. DISCUSSION

The results reported in this rather short chapter demonstrate that the clone 55 transcript is made by at least one other T-helper cell line besides EL4, and this result suggests that this mRNA is probably expressed in the normal immune system, and is not an artifact of the EL4 cell line. The cell line where the transcript is found (CH2.4) was derived from CBA/J (H-2^k) splenocytes, grown in the presence of purified IL2 and irradiated allogeneic lymphocytes (C57/Bl-6 [H-2^b]). The cell line is absolutely dependent on antigen for continued growth, although it will proliferate for a short time in the presence of IL2 alone (C. Havele, personal communication). The growth requirements of the CH2.4 cell line suggest that it is very nearly physiological in its behavior, and reinforces the idea that these cells represent a good model of a 'normal' T-helper cell. What is particularly interesting is that the CH2.4 cells do not express the clone 55 transcript until they are 'activated' by treatment with mitogens, an expression pattern we would expect for lymphokines in normal T-helper cells. Furthermore, the 55 transcript is blocked in the presence of CsA, a result which is also seen for lymphokines expressed in normal activated T-helper cells. This is in direct contrast to the effect of CsA on PMA-treated EL4 cells, where the 55 transcript is 'super-induced.' Further investigations as to the reason for the differences in expression pattern in these

TABLE 10

BIOASSAYS IN WHICH COS CELL SUPERNATANTS PRODUCED FROM CLONE 55 HAVE BEEN TESTED

Lymphokine or Cytokine activity tested	Collaborator
-interleukin 2 -differention of CTLs -growth of primary cultures of splenocytes	C. Havele, University of Alberta, Edmonton
-growth of primary bone marrow cultures -differentiation of a lymphoid precursor cell line	C. Eaves, B.C. Cancer Research Laboratory, Vancouver
-B cell growth factor 1 -B cell differentiation factors -factors which cause B cell immunoglobulin isotype switching (mu to gamma and alpha)	D. Kunimoto, National Institute of Health, Bethesda
-interleukin 3 -GM- CSF -interleukin 4 -erythropoietin -CSF-1	L. Gilbert, University of Alberta, Edmonton
-CTL generation	J. Plate, Rush Medical College, Chicago
-Mast cell growth -Mast cell differentiation	D. Harnish, McMaster University, Hamilton
-macrophage migration -macrophage growth	A. Greenberg, Manitoba Cancer Research Inst., Winnipeg
-supression of B-cell responses	D. Green, University of Alberta, Edmonton

two cell lines may lead to some insight into the mechanisms which control the expression of the clone 55 transcript.

A number of cell lines tested did not express the clone 55 transcript, even after they had been activated by mitogens (Table 8). From the small sample of cell lines surveyed it is not possible to make any firm conclusions regarding the phenotype of the cells which express the clone 55 mRNA. The two cell lines which do make the 55 transcript are both helper T-cell lines, and it is possible that all cells which express the 55 gene are of this phenotype. This suggestion is consistent with the idea that the clone 55 product is a lymphokine-like molecule. The T-cell line BW5147 does not make the 55 transcript, although the cell line has many features to suggest that it is also in the T-helper cell class. It is very likely that even if the 55 transcript is expressed only in T-helper cells, it will not be found in all cells of this sub-class.

From the bioassay data collected to date it is not possible to determine the biological effect of the clone 55 protein product, or even if this protein has any lymphokine or immunoregulatory activity. When the relevant cDNA sequence is inserted, the SV40 based vector and COS cell transient expression system produce both human and mouse IL2, as well as mouse GM-CSF (all active at dilutions greater than 100-fold), and it is very likely that the system also expresses the clone 55 protein. In the face of negative results it is impossible to determine if the clone 55 protein simply has no affect on the bioassays in which it was tested, or if there was insufficient protein present to cause a detectable effect. The COS system produces extremely small quantities of protein, and although these amounts are sufficient to give a positive bioassay for IL2 or GM-CSF, this may not be the case for the protein in question. Future experiments will involve both using procaryotic expression systems to produce large quantities of the protein, and repeating the bioassays at higher concentrations of the clone 55 protein.

Alternatively, failure to find a biological activity in the clone 55 COS cell supernatants may be due to the fact that the protein has not been tried in the 'right' bioassay; perhaps the 'right' bioassay does not yet exist. Several powerful systems have been developed for isolating cDNAs on the basis of expression, and recently some of these have been used to clone the 'interleukin 4' cDNA (Lee *et al.*, 1986; Noma *et al.*, 1986). It is very likely that these systems can and will be applied to clone all of the lymphokine cDNAs for which a well-defined bioassay exists. The cloning of additional lymphokines will then depend on the development of new bioassays which detect novel activities. If a cell secretes many lymphokines, some of the minor lymphokine activities may remain masked by those which predominate, and the minor activities will remain undiscovered unless they can be isolated. Isolating and expressing potentially interesting cDNAs, such as those which are induced during mitogen activation of T-helper cells, would provide a viable method of isolating the corresponding proteins. The task of developing a bioassay to discover what an unknown, potential lymphokine does may therefore become rather more common in the future. These types of investigations may lead to significant insights into lymphocyte biology and the cellular interactions which take place in the immune system. In any case, solving the problem of the function of the clone 55 protein, if in fact it does represent a lymphokine, may require a significant amount of time and imagination.

CHAPTER VIII

FURTHER CHARACTERIZATION OF THE SHORT PMA-INDUCIBLE MMTV TRANSCRIPT FROM EL4

A. INTRODUCTION

The majority of the 136 PMA-induced cDNA clones isolated from EL4 could be grouped into a single large family, and the corresponding transcripts were derived from the retrovirus mouse mammary tumor virus (MMTV) (Chapter VI). Kwon and Weissman (1984) found a similar result, and it would be useful to compare these authors' findings with the results of this study and with the results of more recently published work.

Kwon and Weissman (1984) found that a 24S MMTV-related transcript was induced by PMA in EL4 cells, and reported an 1800 bp cDNA sequence which was derived from the *env* and U3 LTR region of mouse mammary tumor virus (MMTV). The sequence demonstrated a unique 491 bp deletion in the LTR region, and displayed several large open reading frames. Racevskis (1986) showed that exposure of EL4 cells to PMA induced the synthesis of protein products encoded by the MMTV LTR region, and that all glycosylated proteins were derived from one initial nonglycosylated translation product of 21,000 M_r. Furthermore, he reported that PMA treatment of EL4 cells induced a 1-kilobase MMTV transcript, which had not been observed by Kwon and Weissman. The relationship of this shorter induced transcript to Kwon and Weissman's cDNA sequence, as well as the relationship of both of these to the induced LTR protein was unclear. In this chapter we confirm that the predominant PMA-induced transcript in EL4 is about 1-kb, and we report the cDNA sequence of this mRNA. The sequence has a single large open reading frame which begins at the 5'-most ATG and has the potential to code for a protein of molecular weight 22,800.

Transcription of the MMTV provirus has been shown under various conditions to yield three possible transcripts; a full length mRNA and two spliced species (Van Ooyen *et al.*, 1983; Wheeler *et al.*, 1983) (see bottom of Figure 33). The longer spliced transcript codes for the *env* proteins, and the shorter has a single large open reading frame which extends into the 3'LTR and codes for the so-called LTR protein, the function of which remains unknown. All three transcripts initiate within the 5'LTR, downstream of the same promoter-enhancer elements. The relative abundance of the three transcripts is therefore controlled at the level of splicing, and the absolute levels of transcription are largely due to controlling elements in the 5' LTR. The site of initiation of the 1 kb MMTV transcript from EL4 is unclear, yet this information would be useful in defining cis-acting elements which might be regulating the expression of this interesting PMA-inducible mRNA. In this chapter we show that the short PMA-inducible transcript is novel in that it initiates within the *env* region of the provirus.

Additional northern analysis confirms that this short retroviral transcript is specifically blocked in the presence of the immunosuppressive agent CsA, and this expression pattern is the same as that of IL2. This novel PMA-inducible transcription unit in the *env* region of MMTV may therefore provide a model system for studying the regulation of lymphokine gene expression in helper T-cells.

B. RESULTS

1. *Sequence of the Most Abundant PMA-Induced Transcript in EL4.*

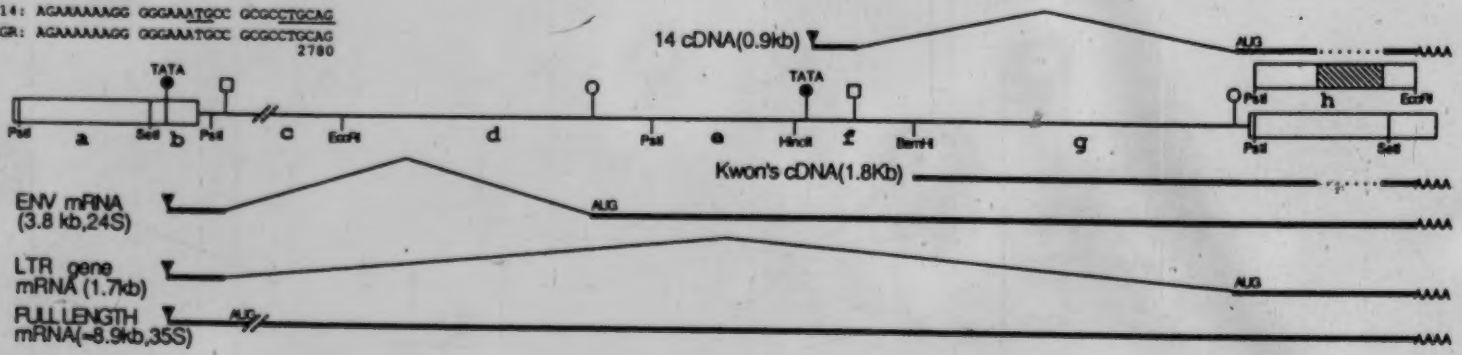
The largest family of induced clones contained 99 members out of the 136, and was represented in 15 of the 20 randomly chosen phage clones which were originally subcloned into pUC13. cDNA inserts from the fifteen members of this family which were subcloned all gave almost identical results when used to probe northern blots of

FIGURE 33. cDNA sequence of the predominant PMA-induced transcript in EL4. The prototype sequence (**14**) is shown at the top of the figure. This has been aligned with the published sequence of the *env* and 3'LTR region for the GR strain of MMTV, using the base numbering of Redmond and Dickson (1983). Differences between the two sequences are marked by an asterisk. Relevant regions underlined in the GR sequence include a TATA promoter consensus sequence at 1408, a splice donor consensus sequence at 1531, a splice acceptor consensus sequence at 2684, and a 10 base pair region which is homologous to a possible control sequence for mouse IL2 at 1569 (see discussion). Relevant regions underlined in the 14 sequence include the *Bgl* II site at 2725, the 5'-most ATG at 2766 and the *Pst* I site at 2775. Only the 5' end of the 14 sequence has been given in detail, since the sequence downstream to the *Pst* I site is nearly identical to that published by Kwon and Weissman (1984). The lower part of the figure gives a schematic drawing of the entire MMTV provirus, and illustrates where clone 14 (above) and the various published transcripts (below) arise (Van Ooyen *et al.*, 1983; Wheeler *et al.*, 1983). The provirus is bounded by two LTR's (open regions); also illustrated are relevant promotor consensus regions (TATA), splice donor consensus regions (open squares), and splice acceptor consensus regions (open circles). The provirus can be divided into a number of fragments, labelled a) through g), by the restriction enzymes indicated. The *Hinc* II site would be at b.p. 1363 of the GR sequence shown above. The various transcripts are indicated by heavy lines, and splice junctions are connected by lighter inverted V's. The presumptive capped 5' ends of the transcripts are indicated by solid inverted triangles, and the 3' ends by poly A tails; the 5' most AUG is also indicated for each mRNA. The lighter dotted lines which interrupt the 3' end of both the clone 14 transcript and Kwon's transcript represent a 491 bp deletion from the proviral LTR sequence which is unique to EL4. Fragment h) was derived from the clone 14 cDNA and also contains this deletion (cross-hatched area). Note that this map is not drawn to scale.

```
14:                                                                                                    GA CTGCCAGATC GCCTTTAAGA AGGACGCCTT CTGGGAAGGA GACGAGTCTG CTCCTCCACG GTGGTTGCCT
GR: CTTTGGGGAA CTCAGGGGCC AATACAAAAC TGGTCCCTAT AAAAAAGAAG TTGCCCCCCA AATATCCTCA CTGCCAGATC GCCTTTAAGA AGGACGCCTT CTGGGAGGGA GACGAGTCTG CTCCTCCACG GTGGTTGCCT
         1380                                                                                                                                                1500

14: TGCGCCTTCC CTGACCAGGG C----------------------probable splice junction----------------------TGCCTTGC GAAGAGCCTT GACCAAGTGC AGTCAGATCT TAACGTGCTT CTTTTAAAAA
GR: TGCGCCTTCC CTGACAAGGG GGTGAGTTTT TCTCCAAAAG GGGCCCTTGG GTTACTTTGG GATTTCTC--///--GTTTTCC AGTGCCTTGC GAAGAGCCTT GACCAAGTGC AGTCAGATCT TAACGTGCTT CTTTTAAAAA
                                                                  1570                  2690

14: AGAAAAAAGG GGGAAATGCC GCGCCTGCAG
GR: AGAAAAAAGG GGGAAATGCC GCGCCTGCAG
                          2780
```



RNA from EL4.E1 cells (3 such northern blots are shown in Figure 17). When the 15 clones were cleaved with the restriction enzyme *Hinf* I, 3 different maps were obtained (data not shown). The three maps differed only in the absence of a *Hinf* I site at one or another of two locations, and were otherwise identical. cDNA inserts representative of the three different *Hinf* I patterns were subcloned into M13mp18 and sequenced. The sequences of all three were nearly identical, but single base changes accounted for the *Hinf* I polymorphisms observed. The prototype of these sequences (clone 14) is shown at the top of Figure 33, where it is aligned with the *env* and 3'LTR region of the MMTV provirus (only the 5' part of the clone 14 sequence is given, since the rest of the sequence was virtually identical to that published by Kwon and Weissman [1984]). The sequences for two of the clones began at the same place (indicated by 14), and the third began 14 bp further upstream. When aligned with the provirus, all three cDNA sequences apparently contained an identical deletion of 1161 bases near their 5' ends (top of Figure 33). All the sequences also contained a 491 bp deletion of the 3'LTR region. The total length of the clone 14 cDNA sequence was 890 bp (excluding any poly A tail).

2. *Expression Pattern of MMTV Related Transcripts in EL4.*

When clone 14 cDNA was used to probe a northern blot from EL4.E1 a number of poly A^+ species were observed (Figure 34). The two larger species represent the full length (circle) and *env* (square) retroviral transcripts, and these are only slightly induced following PMA treatment (compare lanes o] and p]). In contrast, a very short transcript (diamond) is completely absent in control cells, but induced to very high levels following PMA stimulation. Most significantly, this same transcript is almost absent when EL4 cells are induced with PMA in the presence of CsA (lane c]). Using single stranded DNA markers, the size of this highly abundant, short PMA-induced transcript was determined to be ≈920 bases (data not shown). A minor PMA induced

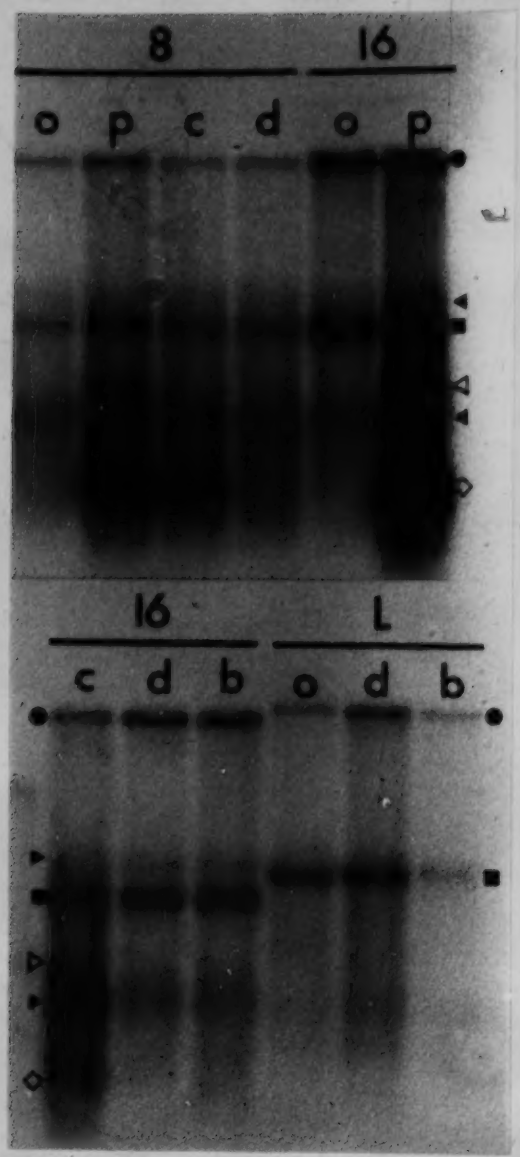


FIGURE 34. Northern blot analysis of poly A+ RNA from EL4, probed with clone 14. RNA was harvested from EL4 cells following either 8 or 16 hours of culture. Additions to the cultures at time=0 were as follows: o) medium alone (RHFM); p) RHFM plus PMA (20ng/ml.); c) RHFM, PMA (20ng/ml.), and CsA (100ng/ml.); d) RHFM plus dexamethasone (1μM.); b) RHFM, dexamethasone (1μM.), and CsA(100ng./ml.). The various bands on the northern are labelled as follows: 1) 18S and 28S ribosomal RNAs (solid arrowheads); 2) the Full Length (circle) and *env* (square) mRNA's of MMTV; 3) a minor PMA-induced transcript (open arrowhead); and 4) the predominant PMA-induced transcript (diamond). The final three lanes (marked L) contain poly A^+ RNA from L1210 cells harvested at 16 hours and treated as outlined above. Separate experiments showed that PMA had no effect on any of the MMTV related transcripts made by L1210 cells.

transcript is also seen at about 22S (open arrowhead), and it is also blocked by CsA. Dexamethasone alone had no effect on the levels of any of the poly A^+ RNAs (lane d]).

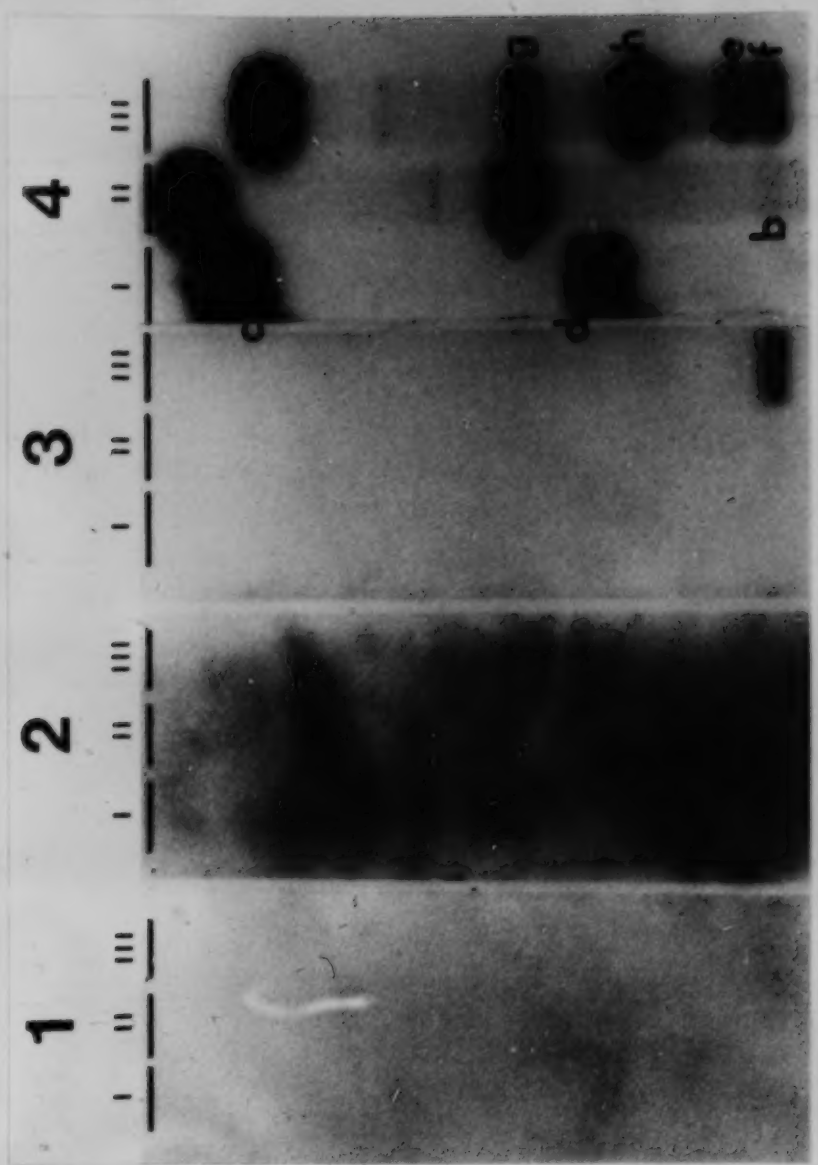
3. *Determining the Initiation Site of the Short MMTV Transcript Using Cap Labelled RNA.*

To investigate where in the MMTV proviral sequence the short PMA-inducible transcript initiates, the technique of cap labelling was used. The 5' cap structure is chemically removed from poly A^+ RNA and specifically replaced with [alpha^{32}P]-GTP. This material will hybridize to all DNA fragments of the provirus which have sequence in common with the viral mRNAs, but if the blot is secondarily treated with RNase A, all single stranded RNA will be hydrolyzed. Only the DNA fragments of the provirus which bracket the transcription start sites will bind the RNA sufficiently close to the CAP structure to prevent hydrolysis of the label, and thus they will remain as labelled bands. The predominant band observed (Figure 35, panel 3) corresponds to fragment f), which is derived from the *env* gene of the MMTV provirus (Figure 33), and this confirms that many retroviral transcripts initiate within this region. Figure 35 also demonstrates that the transcripts initiating within fragment f) are present at high levels in EL4.E1 cells treated with PMA (panel 3), but not in control cultures (panel 2, band f] absent); and in the presence of PMA and CsA these transcripts are reduced below detectable levels (panel 1, band f] absent). The longer MMTV transcripts which initiate within fragment b) are also below detectable levels in these exposures.

4. *Genomic Structure of the MMTV Proviruses Present in EL4.E1.*

MMTV is amplified in EL4 (Dudley and Risser, 1984; Kwon and Weissman, 1984; Racevskis, 1986), and the hypothesis that the short PMA-inducible transcripts arise from a large number of these proviruses led us to investigate specifically the

FIGURE 35. Cap labeled RNA demonstrates that the short MMTV transcripts initiate within the *env* region of the retrovirus. Identical nylon filters containing a series of restriction fragments which span the entire MMTV genome were hybridized to cap-labeled poly A^+ RNA obtained from EL*4* cells cultured for 16 hours under the following conditions: 1) PMA (20 ng./ml.) and CsA (100 ng./ml.) in RHFM; 2) RHFM alone; and 3) PMA (20ng./ml.) in RHFM. Blot 4) was hybridized to a mixture of all of the restriction fragments and was used to locate the various bands, labeled a) through h), which correspond to the segments of the MMTV provirus shown schematically in figure 33. The sizes of the various fragments are as follows: a) 1075 bp., b) 375 bp., c) ≈4000 bp., d) 935 bp., e) 435 bp., f) 365 bp., g) 1045 bp., h) 670 bp. DNA representative of the entire MMTV genome was contained in three plasmids, which were cut with a variety of restriction enzymes to yield the bands seen on blot 4). The plasmids and enzymes used are as follows: lane I, ATCC No. 45005 (*Pst* I gag-pol insert in pBR322) cut with *Pst* I and *Eco* R1; lane II, ATCC No. 45004 (*Pst* I LTR insert in pBR322) cut with *Pst* I and *Sst* I; lane III, pJE1 cut with *Pst* I , *Bam* H1, *Eco* R1, and *Hinc* II. pJE1 was constructed by ligating the *Pst* I -*Bgl* II fragment from ATCC No. 45006 (*Pst* I *env* insert) to the *Bgl* II-*Eco* R1 fragment from clone 14, and placing these in *Pst* I -*Eco* R1 cut pGEM-2(Promega). This yields a composite of the fragments e), f), g), and h) seen in figure 33. The various DNA fragments of the MMTV genome were separated on a 1.5% agarose gel prior to blotting onto a nylon membrane, and the four identical filters were cut from a single blot (12 tracks, repeated in groups of three). Although band b) is fainter than band f) in blot 4), both bands were of equal intensity in the stained agarose gel prior to transfer, and the difference in intensities reflects differences in the relative amounts of probe specific for each fragment. Darker bands above band c), below d), above a) and above g) represent plasmid vector DNA. This experiment has been repeated three times with identical results.


1
2
3
4
a
b
c
d

internal structure of the majority of the (amplified) copies of MMTV, using *Pst* I cut southern blots of EL4.E1 DNA (Figure 36). In the majority of the proviral copies the *Pst* I fragment spanning c)+d) (Figure 33) is intact (expected size ~5000 bp), as is the *Pst* I fragment spanning e)+f)+g) (expected size 1845 bp). On digesting with *Pst* I, the usual MMTV 5' LTR should yield a fragment of 1450 bp (or possibly two fragments of 919 and 536 bp if an internal *Pst* I site is present [Kennedy *et al.*, 1982; Peterson *et al.*, 1985]). A single dark band at 960 bp (Figure 36, lane a]) suggests that a ~490 bp deletion is present in the 5'LTRs of the majority of the proviral copies. The 3'LTR contains only one *Pst* I site, and the numerous lighter bands in lanes a) and 14) represent DNA fragments which contain the 3'LTR plus adjacent cellular sequences. The size of these fragments depends on the location of the nearest *Pst* I site in the flanking DNA, and the fragments will be unique for each different integration site. The number of lighter bands gives an estimate of the number of MMTV proviruses which are present in EL4.E1, about 20 in this case. The idea that the majority of transcribed (but basically intact) proviruses in EL4.E1 have a deletion of ~490 bp in their 3'LTRs is reflected in shorter *env* and full length MMTV transcripts. The two transcripts from EL4 are both about 500 bases shorter than those from L1210 cells, which do not have this deletion (Figure 34).

C. DISCUSSION

In this study we have used subtractive hybridization to obtain a large number of cDNA clones which collectively yield a representative sample of all the transcripts which are induced in the T-cell line EL4.E1 following PMA induced 'differentiation.' Fully 75% of these induced clones are derived from the *env* and 3'LTR region of MMTV. When aligned with the provirus, the DNA sequences of 3 representative MMTV related cDNA clones show a number of interesting features. All three have an

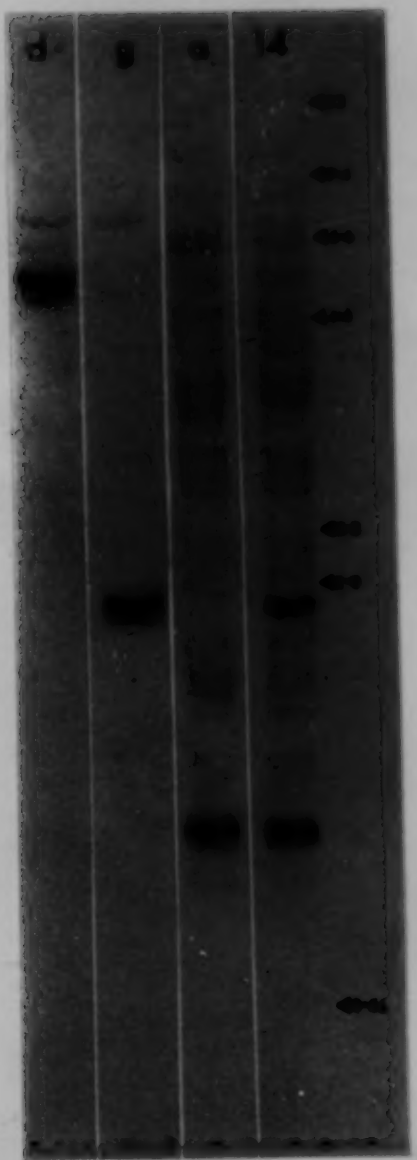


FIGURE 36. Southern blot analysis of EL4 genomic DNA cut with *Pst* I and probed with various segments of the MMTV provirus. The probes d), g), and a) correspond to those segments of MMTV illustrated in figure 33; probe 14 is the entire cDNA insert from clone 14. The location of marker DNA (*Hind* III cut lambda) is shown at the left, with fragment sizes given in kb.'s. The measured size of the darkest bands in each of the various blots is as follows: d), ~5200 bp.; g), 1830 bp.; a), 960 bp.; 14), 1830 and 960 bp.

identical 491 bp deletion in the 3'LTR region, which appears to be a true deletion, is unique to EL4, and has been described previously by Kwon and Weissman (1984). All three cDNA sequences also have an identical 1161 b.p. deletion near their 5' ends, which has not been described previously. This 'deletion' likely represents a splicing event in the original retroviral transcript, since the proviral sequence contains appropriately placed splice donor and acceptor consensus regions (underlined in the GR sequence at the top of Figure 33). As well, in the majority of proviral sequences present in EL4.E1, the 1845 b.p. *Pst* I fragment spanning the *env* region is intact as determined by southern blotting, and there is no indication of a 684 b.p. *Pst* I fragment, which we would expect if the deletion were present in one of the genomic copies of the provirus itself (Figure 36 lanes g] and 14]). Finally, all three of the cDNA sequences begin at nearly the same place, within the *env* region of the provirus, and none of them contains non-viral sequences at the 5' end, which would be likely if a cellular promoter and control region were fused upstream of the *env* region of an incomplete provirus.

The fact that the three cDNA sequences described above all begin at nearly the same place on the proviral sequence may have been purely fortuitous. cDNA clones often do not extend to the very 5' ends of the mRNAs from which they were originally copied. S1 mapping would not rule out the possibility that the transcripts began at the usual site in the 5' LTR, copied just a few bases, and were finally spliced to a region just upsteam of the clone 14 sequence. To prove that the short PMA-induced transcripts do initiate within the *env* region we used the technique of cap labelling (Figure 35). When DNA fragments representative of the entire MMTV genome were probed with cap labelled poly A^+ RNA from EL4 cells treated with PMA, a large number of MMTV transcripts were demonstrated to initiate within fragment f) (Figure 35). These transcripts were not present in corresponding cap labelled RNA from control cells, and

were markedly reduced in RNA from cells treated with PMA and CsA. When considered together with the results of northern analysis (Figure 34), these observations present a coherent picture of one or a few short MMTV transcripts which are induced in EL4.E1 cells in the presence of PMA and blocked in the presence of PMA and CsA. Unlike any previously described retroviral mRNAs (which all initiate within the 5'LTR), these inducible transcripts initiate within the *env* gene of the provirus.

The observation that the short PMA-inducible MMTV transcripts are actually initiating within the *env* gene of the provirus led us to search for possible promoters in the proviral sequence itself. The promoter consensus sequence TATAAAA is found four times in all of the available MMTV sequences (Bionet data base and recent literature): once in each LTR (25 b.p. upstream of the known transcription start site); once within the coding region of the *pol* gene (Deen and Sweet, 1986); and once in the *env* region, 24 b.p. upstream of the start of sequence 14 (underlined in the GR sequence at the top of Figure 33). Although the entire MMTV sequence is not available (the 3' region of the *gag* gene has not been sequenced), the fact that this promoter consensus sequence occurs infrequently over ~7 Kb. of sequence, combined with the fact that it occurs just upstream of the clone 14 sequence (at the correct distance) makes it likely that many of the short PMA-inducible transcripts are generated from this internal viral promoter. Sequence analysis revealed that one of the three cDNA clones began 14 bp further upstream than the clone 14 sequence shown in Figure 33. Examination of the proviral sequence reveals a second near-consensus promoter sequence (TACAAAA) which is 15 bp 5' to the 'TATA box' underlined in the GR sequence (Figure 33). This second promoter (or a mutated form of it) may be responsible for generating some of the PMA-induced MMTV transcripts.

Kwon and Weissman (1984) were the first to demonstrate that in EL4 cells PMA induces a transcript derived from the *env* and 3'LTR region of MMTV. The size of

their transcript (about 24S) is considerably greater than the ≈920 base transcript reported here, and their sequence extends continuously through the *env* gene from the *Bam* HI site, and shows no apparent splice in the 5' region (bottom of Figure 33), although both transcripts have the identical 491 deletion in the 3'LTR region. Since Kwon and Weissman were cloning and subcloning into a *Bam* HI site, it is possible that their sequence may not extend to the 5' end of their 24S transcript. The difference in size between the two inducible transcripts might be accounted for if the 24S transcript (which for some reason predominates in Kwon and Weissman's sub-clone of EL4) represents the unspliced version of the shorter PMA-inducible transcript. In this way both transcripts might be regulated by the same PMA-inducible promoter within the *env* gene of the provirus. Our northern blots show a small amount of a longer PMA-induced transcript which probably corresponds to the unspliced form (Figure 34, open arrowhead). Alternatively, Kwon and Weissman's sequence may simply represent the 3' portion of an *env* transcript, which initiates from the 5'LTR promoter and is slightly induced in our cells.

Our results support those of Racevskis (1986) who probed northern blots with the MMTV LTR sequence and showed that the predominant PMA-inducible transcript in EL4 cells is ≈1 kb in length. In the same report Racevskis demonstrated that exposure of EL4 cells to PMA induced the synthesis of protein products encoded by the MMTV LTR region, and that all glycosylated proteins were derived from one initial nonglycosylated translation product of 21,000 M_r. The clone 14 cDNA sequence, which we believe is representative of the short PMA-induced transcripts, has a single large open reading frame which begins at the 5'-most ATG (underlined in Figure 33) and has the potential to code for a protein of molecular weight 22,800. It is very likely, therefore, that the short PMA-induced transcripts give rise to the PMA-induced protein reported by Racevskis. This is exactly the same protein which would be potentially

coded for by the 'LTR gene' transcript (providing the 491 bp deletion had occurred in the 3'LTR), since both this transcript and the clone 14 transcript use the same splice acceptor site (bottom of Figure 33). However, the cap-labelling data demonstrate that the majority of PMA-induced transcripts do not initiate in the 5'LTR, as would be expected for the 'LTR gene' transcripts.

What is the (genomic) structure of the MMTV provirus in EL4 from which the short PMA-inducible transcripts arise? Clearly the provirus has a 491 bp deletion in the U3 region of the 3' LTR. We have suggested that the transcripts initiate from a viral promoter, and are spliced, and this would imply that the responsible provirus is intact in the *env* region (fragments f] and g] in Figure 33). A number of authors (Dudley and Risser, 1984; Kwon and Weissman, 1984; Racevskis, 1986) have shown that the MMTV provirus is amplified in EL4 (about 20 copies, see also Figure 36), and since many of the proviral copies have the 3'LTR deletion (Racevskis, 1986) it would in theory be difficult to determine just which proviral copy gives rise to the short PMA-inducible transcripts.

An alternate hypothesis is that the short PMA-inducible transcripts arise from a large number of the MMTV proviruses which are present in EL4. Although by no means proven, this hypothesis is supported by two observations. Relative to IL2 a very large number of PMA-inducible MMTV transcripts are made. Since the IL2 and MMTV transcripts are about the same size, we can make an estimate of the relative abundance of these two mRNAs by comparing their relative frequencies in the original 136 induced clones (see Chapter VI). The MMTV transcript is about 25X as abundant as the IL2 transcript, since 99 of the 136 were in the 'clone 14' family, whereas 4 of the 136 clones were IL2. Furthermore, significant polymorphism was seen in the three MMTV cDNA clones which were sequenced; these were primarily single base changes, but were more frequent (≈ 1 change per 60 bases) than would be expected simply from

copying errors caused by reverse transcriptase during the first step of the cDNA cloning (estimated to be 1 error per 600 bp [Battula and Loeb, 1974]). A simple explanation for the large number of MMTV transcripts produced, and the apparent polymorphism between the transcripts, is that these transcripts are being copied from a large number of proviral genes which are nearly identical, but have diverged since (or possibly during) their original amplification. Given this hypothesis it is reasonable to determine the overall genomic structure of the majority of the MMTV proviruse present in EL4.E1.

The internal structure of the majority of the amplified copies of MMTV was determined using *Pst* I cut southern blots of EL4 DNA. These data, together with *Msp* I cut southern blots probed with the LTR (Racevskis, 1986), demonstrate that the majority of MMTV proviruses present in EL4 appear to be intact in the *gag*, *pol*, and *env* regions, but contain ≈490 bp deletions in both LTRs. Any (or perhaps all) of these proviruses could in theory give rise to the short PMA-inducible transcripts. At least some of these appear to be transcriptionally active, giving rise to the *env* and full length MMTV transcripts present in EL4 (which are shorter than expected because of the deletion in the 3' LTR; Figure 34 and Results).

Our findings are very similar to those of Michalides *et al.* (1985), who found MMTV to be amplified in a number of T-cell leukemias from GR mice. All copies of the amplified provirus appeared to have a similar sized deletion which was present in both LTRs. The size and location of the 3'LTR-deletion was unique for each particular leukemia, although the same general area of the LTR was affected in each case. In Figure 37 the size and location of the various 3'LTR deletions described by these authors is compared to that found in EL4. Whether or not the 3'LTR deletion facilitates transcription of the short PMA-inducible MMTV transcript in EL4 cells is unknown. Michalides *et al.* (1985) apparently did not look for PMA or mitogen-induced MMTV transcripts in their T-cell lines. Wellinger *et al.* (1986) have also found alterations in a

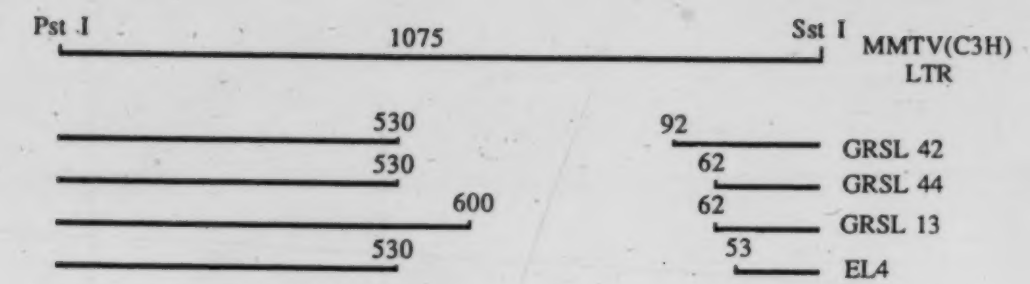


FIGURE 37. Deletions found in the MMTV LTR for various murine T-cell leukemias. At the top of the figure is shown the LTR region of MMTV corresponding to fragment a) in Figure 33, and the length of the fragment in bp is given above it. Below are shown the location and size of various LTR deletions which have been described, where the numbers give the distance from the Pst I site to the deleted segment (left), and the distance from the Sst I site to the deleted segment (right). The GRSL cell lines were described by Michalides et al. (1985). The Sst I site is 110 bp 5' to the cap site of the full length, env, and LTR transcripts made by MMTV.

similar region of the LTR in exogenous MMTV proviral copies present in a murine kidney adenocarcinoma cell line. Dekaban and Ball (1984) reported that an MMTV-related retrovirus can cause primary murine thymic lymphomas, but they have not examined the structure of the LTRs in this virus. It is possible that alterations in one or both LTRs may influence the tumorogenic properties of MMTV.

Wheeler *et al.* (1983) reported several short (1-1.2 kb) MMTV transcripts which were present in some virus-induced mouse mammary tumors and in some pre-neoplastic mammary tissue. These transcripts cross-reacted strongly with probes from the *env* region, or to probes from the LTR, but did not cross-react with a 220 bp probe from the 'viral leader sequence' (that segment of the provirus approximately between the 3' end of the left LTR and the first splice donor site; see Figure 33). The authors suggest that these short transcripts probably do not originate from the 5'LTR promoter, and it is possible they initiate in the *env* region of the provirus, downstream of the internal promoter described above.

The longer MMTV transcripts are not induced by dexamethasone in EL4.E1 cells (Figure 34), presumably because the deletion in the 5'LTR (which, if identical to the 3'LTR deletion, removes a segment 5' to base -163) also removes most of the glucocorticoid responsive region. This steroid inducible enhancer region normally acts on the promoter in the 5'LTR (Diggelmann *et al.*, 1985).

In this work we have described a MMTV transcript whose expression exactly parallels that of mouse IL2 in EL4 cells. The transcript is induced by PMA, and this induction is blocked by the immunosuppressive agent CsA. A number of authors have identified DNA sequences present in the 5' region of the IL2 gene which may be important in controlling the transcription of this lymphokine gene. Do similar sequences exist in the *env* region of MMTV which may be regulating transcription of the short MMTV transcript? Stanley *et al.* (1985) described a ten base sequence which

was highly conserved in a number of hemopoietic growth factor genes, including murine GM-CSF and murine IL2. The relevant sequence in mouse IL2 is 5'-GGGATTTCAC-3', and it occurs ~175 bp 5' to the TATA promoter consensus sequence. A computer search of all available MMTV sequences shows that a perfect match with the 10 base mouse IL2 sequence does not occur, and that a sequence with one mis-match occurs only once. This sequence 5'-GGGATTTCTC-3' (the mis-match is underlined) occurs ~160 bases 3' to the putative TATA promotor, in the first intron of the short PMA-inducible transcript (top of Figure 33). In a number of genes cis-acting control sequences have similarily been found within introns (Weinberger *et al.*, 1986).

Fujita *et al.* (1986) described an 18 base consensus sequence which is present twice in the 5' flanking region of the mouse IL2 gene, as well as in the human IL2 gene and in a number of other related genes which are expressed in activated T-cells. The 18 base consensus sequence, along with the corresponding proximal sequence found in the mouse IL2 gene, are shown in Figure 38 (the numbers indicate the distance from the transcription start site). A sequence highly homologous to the mouse IL2 proximal sequence is found in the *env* region of the MMTV provirus, and this is shown below the mouse sequence. The MMTV sequence is inverted with respect to the IL2 sequence (note that the numbers indicating the position of the two sequences have the opposite progression), but this is not unreasonable, since the control sequences in the 5' flanking region of the mouse IL2 gene apparently function in either orientation (Fujita *et al.* [1986]).

Further experiments will be required to determine whether or not the two sequences described here actually do play a role in regulating the expression of the short PMA-inducible MMTV transcript in EL4 cells. In any case, given the similarities in expression patterns, it is likely that similar factors act to regulate the expression of both

CONSENSUS		AGAAAGGAGGAAAAACTG
Murine IL2 (proximal)	-145	AGAA-G-AGGAAAAACAA -130
MMTV (*env*)	-195	AGGA-GAAGGAAAAGCAA -211

FIGURE 38. Homologous nucleotide sequence found in the 5' flanking region of mouse IL2 and the *env* region of MMTV. The numbers indicate the position of the sequences relative to the transcription start site(+1) for these two PMA inducible mRNA's from EL4. The consensus sequence was derived from comparison of a large number of similar sequences found in the 5' flanking region of lymphokine genes (Fujita *et al.*, 1986), and gaps (-) have been introduced to maximize homology between the three sequences. Nucleotides which differ from those found in the consensus sequence have been underlined. There are two such sequences in the 5' flanking region of the mouse IL2 gene; only the proximal one is shown here, since it is most homologous to the sequence found in the MMTV *env* region.

the IL2 gene and the short MMTV transcript. For this reason we believe that the PMA-inducible transcription unit in the *env* region of MMTV will serve as a useful model system for studying the regulation of lymphokine gene expression in T-helper cells.

CHAPTER IX

GENERAL DISCUSSION AND CONCLUSIONS

The objective of this study was to isolate and characterize a number of the mRNAs which are induced in the mouse T-lympohoma cell line EL4 when it is treated with PMA. The results presented show that established methods of subtractive hybridization can be used to isolate cDNA clones which represent such PMA-induced mRNAs. This method appears to be very reliable, since of the 136 independent clones chosen, only 2 were false positives (*i.e.* they contained sequences which were not induced by PMA). The method of subtractive hybridization was used because it is the method of choice for isolating differential genes representing mRNAs from the low abundance class of transcripts (Chapter I). Of the PMA-induced cDNAs isolated in this study, it is unlikely that any of them represent low abundance transcripts (judging by the strength of the signal detected on northern blots), although this has not been proven rigorously. It is possible that if more clones had been examined, some which contained sequences representing low abundance (PMA-induced) transcripts would have been found.

The technique of subtractive hybridization (particularly when used to generate subtracted cDNA libriaries) requires relatively large quantities of poly A+ RNA, and tumor cell lines or well established hybridomas are the only clonal cell populations (at least in higher eucaryotes) which can easily fulfill these requirements. This places a limitation on the biological systems to which subtractive hybridization can be applied if a strictly defined set of differential genes is to be cloned. This represents a serious drawback of the method, as was demonstrated in this report. The PMA-induced differentiation of EL4 was used as a model for the differentiation events which occur

when normal T-helper lymphocytes encounter antigen, become activated, and begin to secrete lymphokines. However, fully 75% of the PMA-induced transcripts in EL4 are derived from the retrovirus MMTV, and this is most probably due to the transformed state of the cell line, and does not reflect what happens in a normal T-helper cell. Thus the very property of EL4 which allows one to obtain with ease large quantities of mRNA also introduced a huge artifact into the biological differential under examination. Similar results might be expected whenever tumor cell lines are used as models for biological differentials.

Despite the fact that the majority of PMA-induced transcripts in EL4 are derived from MMTV, the cell does make a number of transcripts which are also induced in normal T-helper cells (such as IL2 and GM-CSF, Chapter III). The presence of the MMTV transcripts is undesirable largely because they make it very difficult to find clones representing these more 'normal' PMA-induced transcripts. As was shown in this study, if a large enough sample of clones is surveyed, it is possible to isolate clones containing more 'normal' cellular sequences such as IL2 and clone 55 (which was shown to be induced by mitogens in another helper T-cell line). A method of eliminating the MMTV sequences from the subtracted library and/or probe would simplify considerably the task of isolating additional PMA-induced clones containing 'normal' cellular sequences. This idea is generally applicable to any situation where a differential set of genes is being cloned by using subtractive hybridization. The highest abundance differential genes are most easily cloned, but unless they can then be eliminated from the selection process the isolation of lower abundance differential genes becomes progressively more difficult.

Although the PMA-induced MMTV transcript in EL4 makes the isolation of other sequences more difficult, the fact that this retroviral transcript is induced is in itself interesting. Inducible viral genes have historically provided useful model systems for

studying the processes which control transcription of more normal eucaryotic genes. The results in Chapter XIII demonstrate that the PMA-induced MMTV transcript is blocked in the presence of CsA, and this suggests that this retroviral transcript may provide a very good model system for studying the control of lymphokine gene expression (since the expression of known lymphokine genes such as IL2 and GM-CSF is also blocked in the presence of CsA, Chapter III). The cis-acting sequences which control the expression of the PMA-inducible MMTV transcript are likely in the *env* region of the retrovirus itself, since this is where the transcript initiates.

In terms of expression patterns, the characteristics of the PMA-induced sequences described in this study are surprisingly diverse. The known lymphokine mRNAs (IL2 and CSF-GM) are blocked by CsA and 'superinduced' by cycloheximide. The MMTV transcript is blocked by CsA and unaffected by cycloheximide, whereas the 'clone 55' transcript (which may code for a lymphokine) is unaffected by cycloheximide, but 'superinduced' by CsA (at least in EL4 cells). The transcripts for carbonic anhydrase II and chondroitin-sulfate core protein are clearly induced by PMA, but are unaffected by any of these other agents. The meaning of these observations with respect to the mechanisms which control the expression of the various transcripts is unknown at present.

As was discussed in Chapter I, the differential approach to cloning genes may well succeed in isolating differentially expressed sequences, but determining the biological significance of the differential may be difficult. The biological significance of the fact that PMA induces the expression of IL2 and GM-CSF in EL4 cells is understandable considering the fact that the cell line mimics a normal T-helper lymphocyte. The biological significance of the fact that PMA induced the expression of carbonic anhydrase is more obscure, although it may be that during PMA activation the

intracellular pH shifts, and induction of this enzyme may be a simple homeostatic response. The biological significance of the observation that chondroitin sulfate is induced by PMA is equally obscure, although it may bear some relationship to the shape and adherence properties of the activated cells, as discussed in Chapter VI. Since the function of the clone 55 transcript is unknown, the biological significance of its induction cannot be considered. Finally, I suggest that there is no biological significance to the fact that the short MMTV transcript is induced by PMA in EL4 cells; it may be a simple coincidence due to the structure of the (amplified) proviral genomes, their location, and the transcription factors which are present in PMA treated EL4 cells. (Alternatively, the short PMA-induced MMTV transcript may have an important function in the biology of the virus itself.) One can also ask if any of these PMA-induced transcripts are induced in normal T-helper cells. In other words, are they biologically significant in terms of T-helper cell function, rather than simple idiosyncrasies of EL4? In the case of IL2, GM-CSF, and the 'clone 55' transcript, the transcripts are induced in more normal mitogen treated T-helper cells, and at least the first two are known to be biologically significant. Whether or not the 'clone 55' transcript is important to the biology of T-helper cells remains to be determined.

When cloning differential genes, the use of a tumor cell line as a model for a biological differential has the advantage that large quantities of RNA can be obtained, so that the construction of subtracted libraries and preparation of subtracted probe is easily accomplished. However, such an approach also has a number of disadvantages, such as artifacts due to the transformed state, and 'model bias' which means that the observations in the tumor model must always be compared to the more normal cells to eliminate idiosyncrasies present in the model. The ideal situation would be one where a specific biological differential could be studied directly by examining small numbers of normal cells. Since quantities of RNA will be limiting in this case, new methods will

be required to clone the differential genes of interest. For example, it may be possible to clone cDNA libraries from both populations of mRNA, and then generate subtracted probes or libraries from these initial libraries. If such a subtractive process were iterative, a general method would then exist to remove high abundance differential sequences from the differential set, and in theory the total set of differential genes could be cloned.

POSTSCRIPT

Since this thesis was completed and submitted to the examination comittee, two important developments have occurred. The first of these was the discovery by D. Green that the COS cell supernatants containing the clone 55 protein have an effect on the development of IgM secreting B cells *in vitro*. The second was the publication by Kinashi *et al.* (*Nature* **324**, 70-73; November 6, 1986) of the cDNA sequence encoding murine T-cell replacing factor (TRF; identical to BCGF II). In fact the published sequence for TRF is identical to the clone 55 cDNA sequence over 1478 bases, with only six base differences. The sequence of the TRF protein (deduced from the cDNA sequence) is identical to the clone 55 protein described in chapter VI, except for a substitution of histidine to arginine at amino acid 79. The significance of this substitution is currently under investigation. It is virtually certain then that the clone 55 cDNA codes for the lymphokine TRF (which will probably be named interleukin 5). Kinashi *et al.* have shown that this lymphokine induces both proliferation and IgM secretion in the B-cell line BCL_1, and that it induces secondary anti-dinitrophenol (DNP) IgG synthesis *in vitro* by DNP-primed B cells.

BIBLIOGRAPHY

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (1983). In *Molecular Biology of the Cell*. (New York: Garland Publishing), p.409.

Alt, F.W., Kellems, R.E., Bertino, J.R., and Schimke, R.T. (1978). Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured murine cells. *J. Biol. Chem.* **253**, 1357-1370.

Alt, F.W., Enea, V., Bothwell, A.L.M., and Baltimore, D. (1979). Probes for specific mRNAs by subtractive hybridization: anomalous expression of immunoglobulin genes. In *Eucaryotic Gene Regulation* , **14**, T. Maniatis, R. Axel, and C.F. Fox, eds. (New York: Academic Press), pp. 407-419.

Atkinson, P.W., King, J.A., and Hynes, M.J. (1985). Identification of the *aciA* gene controlled by the *amdA* regulatory gene in *Aspergillus nidulans* . *Curr. Genet.* **10**, 133-138.

Baxt, W., Yates, J.W., Wallace, H.J., Holland, J.F., and Spiegelman, S. (1973). Leukemia-specific DNA sequences in leukocytes of the leukemic member of identical twins. *Proc. Nat. Acad. Sci. USA* **70**, 2629-2632.

Battula, N. and Loeb, L.A. (1974). The infidelity of avian myeloblastosis virus deoxyribonucleic acid polymerase in polynucleotide replication. *J. Biol. Chem.* **249**, 4086-4093.

Berg, P. (1981). Dissections and reconstruction of genes and chromosomes. *Science* **213**, 296-303.

Berger, S.L. and Birkenmeier, C.S. (1979). Inhibition of intractable nucleases with ribonucleoside-vanadyl complexes: isolation of messenger ribonucleic acid from resting lymphocytes. *Biochemistry* **18**, 5143-5149.

Berlin, V. and Yanofsky, C. (1985). Isolation and characterization of genes differentially expressed during conidiation of *Neurospora crassa* . *Mol. Cell. Biol.* **5**, 849-855.

Biggin, M.D., Gibson, T.J., and Hong, G.F. (1983). Buffer gradient gels and 35S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA*. **80**, 3963-3965.

Bishop, J.O., Morton, J.G., Rosbash, M., and Richardson, M. (1974). Three abundance classes in HeLa cell messenger RNA. *Nature* **250**, 199-204.

Bleackley, R.C., Caplan, B., Havele, C., Ritzel, R.G., Mosmann, T.R., Farrar, J.J., and Paetkau, V. (1981). Translation of lymphocyte mRNA into biologically active interleukin 2 in oocytes. *J. Immunol.* **127**, 2432-2435.

Bleackley, R.C., Havele, C., and Paetkau, V. (1982). Cellular and molecular properties of an antigen-specific cytotoxic T lymphocyte line. *J. Immunol.*

128, 758-767.

Bleackley, R.C., Horak, H., McElhaney, J., Havele, C., Shaw, A.R.E., Turner, A.R., and Paetkau, V.H. (1983). Translation of lymphocyte mRNA for mouse colony stimulating factor. *Nucl. Acids Res.* **11**, 3027-3035.

Bourdon, M.A., Oldberg, A., Pierschbacher, M., and Ruoslahti, E. (1985). Molecular cloning and sequence analysis of a chondroitin sulfate proteoglycan cDNA. *Proc. Natl. Acad. Sci.* **82**, 1321-1325.

Bourdon, M.A., Shiga, M., and Ruoslahti, E. (1986). Identification from cDNA of the precursor form of a chondroitin sulfate proteoglycan core protein. *J. Biol. Chem.* **261**, 12534-12537.

Branch, D.R. and Gilbert, L.J. (1986) Practical colorimetric assay for hematopoietic growth factors. *J. Tissue Culture Methods*, in the press.

Britten, R.J., Graham, D.E., and Neufeld, B.R. (1974). Analysis of repeating DNA sequences by reassociation. *Methods Enzymol.* **29**, 363-406.

Britton, S., and Palacios, R. (1982). Cyclosporin A: usefulness, risks and mechanism of action. *Immunological Reviews* **65**, 5-22.

Bunges, D., Hardt, C., Rollinghoff, M., and Wagner, H. (1981). Cyclosporin A mediates immunosuppression of primary cytotoxic T cell responses by impairing the release of interleukin 1 and interleukin 2. *Eur. J. Immunol.* **11**, 657-661.

Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shimer, S., and Cerami, A. (1986). Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. *Proc. Natl. Acad. Sci.* **83**, 1670-1674.

Chien, Y., Becker, D.M., Lindsten, T., Okamura, M., Cohen, D.I., and Davis, M.M. (1984). A third type of murine T-cell receptor gene. *Nature* **312**, 31-35.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.

Cochran, B.H., Reffel, A.C., and Stiles, C.D. (1983). Molecular cloning of gene sequences regulated by platelet-derived growth factor. *Cell* **33**, 939-947.

Crampton, J., Humphries, S., Woods, D., and Williamson, R. (1980). The isolation of cloned cDNA sequences which are differentially expressed in human lymphocytes and fibroblasts. *Nucl. Acids Res.* **8**, 6007-6017.

Curtis, P.J. (1983). Cloning of mouse carbonic anhydrase mRNA and its induction in mouse erythroleukemic cells. *J. Biol. Chem.* **258**, 4459-4463.

Curtis, P.J., Withers, E., Demuth, D., Watt, R., Venta, P.J., and Tashian, R.E. (1983). The nucleotide sequence and derived amino acid sequence of cDNA coding for mouse carbonic anhydrase II. *Gene* **25**, 325-332.

Davis, M.M., Cohen, D.I., Nielsen, E.A., De Franco, A.L., and Paul, W.E. (1982). The isolation of B and T cell-specific genes. In *B and T cell tumors*, University of California at Los Angeles Symposium, **24**, E. Vitteta and C.F. Fox, eds. (New York: Academic Press), pp. 215-220.

Davis, M.M., Cohen, D.I., Nielsen, E.A., Steinmetz, M., Paul, W.E., and Hood, L. (1984). Cell-type-specific cDNA probes and the murine I region: the localization and orientation of Ad alpha. *Proc. Natl. Acad. Sci. USA* **81**, 2194-2198.

Davis, R.W., Botstein, D., and Roth, J.R. (1980). In *Advanced Bacterial Genetics*, (Cold Spring Harbor: Cold Spring Harbor Laboratories), pp. 106-108.

Deen, K.C. and Sweet, R.W. (1986). Murine mammary tumor virus *pol* -related sequences in human DNA: characterization and sequence comparison with the complete murine mammary tumor virus *pol* gene. *J. Virol.* **57**, 422-432.

Dekaban, G.A., and Ball, J.K. (1984). Integration of type B retroviral DNA in virus-induced primary murine thymic lymphomas. *J. Viro.* **52**, 784-792.

Depper, J.M., Leonard, J.W., Robb, R.J., Waldmann, T.A., and Greene, W.C. (1983). Blockade of the interleukin-2 receptor by anti-tac antibody: inhibition of human lymphocyte activation. *J. Immunol.* **131**, 690-696.

Diggelmann, H., Buetti, E., Owen, D., Fasel, N., Firzlaff, J., and Vessaz, A.L. (1985). A short segment of viral DNA is required for the stimulation of transcription of mouse mammary tumor genes by glucocorticoid hormones. In *Viral Messenger RNA*, Y. Becker, ed. (Boston: Martinus Nijhoff Publishing), pp. 339-353.

Dudley, J. and Risser, R. (1984). Amplification and novel locations of endogenous mouse mammary tumor virus genomes in mouse T-cell lymphomas. *J. Virol.* **49**, 92-101.

Dudley, J.P., Arfsten, A., Hsu, C.L., Kozak, C., and Risser, R. (1986). Molecular cloning and characterization of mouse mammary tumor proviruses from a T-cell lymphoma. *J. Virol.* **57**, 385-388.

Dworkin, M.B., and Dawid, I.B. (1980). Construction of a cloned library of expressed embryonic gene sequences from *Xenopus laevis*. *Develop. Biol*. **76**, 435-448.

Elliott, J.F., Lin, Y., Mizel, S.B., Bleackley, R.C., Harnish, D.G., and Paetkau, V., (1984). Induction of interleukin 2 messenger RNA inhibited by cyclosporin A. *Science* **226**, 1439-1441.

Farrar, J.J., Fuller-Farrar, J., Simon, P.L., Hilbiker, M.L., Stadler, B.M., and Farrar, W.J. (1980). Thymoma production of T cell growth factor (interleukin 2). *J. Immunol.* **125**, 2555-2558.

Farrar, J.J., Benjamin, W.R., Hilfiker, M.L., Howard, M., Farrar, W.L., and Fuller-Farrar, J. (1982). The biochemistry, biology, and role of interleukin 2 in the induction of cytotoxic T cell and antibody forming B cell responses.

Immunological Reviews **63**, 129-166.

Fathman, C.G. and Fitch, F.W. (1984). Long-term culture of immunocompetent cells. In *Fundamental Immunology*, W.E. Paul, ed. (New York: Raven Press), pp. 781-809.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochem.* **132**, 6-13.

Foster, D.N., Schmidt, L.J., Hodgson, C.P., Moses, H.L. and Getz, M.J. (1982). Polyadenylated RNA complementary to a mouse retrovirus-like multigene family is rapidly and specifically induced by epidermal growth factor stimulation of quiescent cells. *Proc. Natl. Acad. Sci.* USA **79**, 7317-7321.

Fotedar, A., Boyer, M., Smart, W., Widtman, J., Fraga, E., and Singh, B. (1985). Fine specificity of antigen recognition by T cell hybridoma clones specific for Poly-18: a synthetic polypeptide antigen of defined sequence and conformation. *J. Immunol.* **135**, 3028-3033.

Freeman, G.J., Clayberger, C., DeKruyff, R., Rosenblum, D.S., and Cantor, H. (1983). Sequential espression of new gene programs in inducer T-cell clones. *Proc. Natl. Acad. Sci. USA* **80**, 4094-4098.

Fujita, T., Shibuya, H., Ohashi, T., Yamanishi, K., and Taniguchi, T. (1986). Regulation of the human interleukin-2 gene: functional DNA sequences in the 5' flanking region for the gene expression in activated T lymphocytes. *Cell* **46**, 401-407.

Fung, M.C., Hapel, A.J., Ymer, S., Cohen, D.R., Johnson, R.M., Campbell, H.D., and Young, I.G. (1984). Molecular cloning of cDNA for murine interleukin-3. *Nature* **307**, 233-237.

Fuse, A., Fujita, T., Yasumitsu, H., Kashima, N., Hasegawa, K., and Taniguchi, T. (1984). Organization and structure of the mouse interleukin-2 gene. *Nucl. Acids Res.* **12**, 9323-9331.

Galau, G.A., Klein, W.H., Davis, M.M., Wold, B.J., Britten, R.J., and Davidson, E.H. (1976). Structural gene sets active in embryos and adult tissues of the sea urchin. *Cell* **7**, 487-505.

Germain, R.N. (1986). The ins and outs of antigen processing and presentation. *Nature* **322**, 687-689.

Gershenfeld, H.K. and Weissman, I.L. (1986). Cloning of a cDNA for T cell-specific serine protease from a cytotoxic T lymphocyte. *Science* **232**, 854-858.

Gillis, S., Ferm, M.M., Ou, W., and Smith, K. (1978). Cell growth factor: parameters of production and a quantitative microassay for activity. *J. Immunol.* **120**, 2027-2032.

Gillis, S. and Watson, J. (1980). Biochemical and biological characterization of

lymphocyte regulatory molecules. *J. Exp. Med.* **152**, 1709-1719.

Gluzman, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182.

Gough, N.M., Gough, J., Metcalf, D., Kelso, A., Grail, D., Nicola, N.A., Burgess, A.W., and Dunn, A.R. (1984). Molecular cloning of cDNA encoding a murine haemopoietic growth regulator, granulocyte-macrophage colony stimulating factor. *Nature* **309**, 763-767.

Goverman, J., Hunkapiller, T., and Hood. L. (1986). A speculative view of the multicomponent nature of T cell antigen recognition. *Cell* **45**, 475-484.

Granelli-Piperno, A., Andrus, L., and Steinman, R.M. (1986). Lymphokine and non-lymphokine mRNA levels in stimulated human T-cells: kinetics, mitogen requirements and effects of cyclosporin A. *J. Exp. Med.* **163**, 922-937.

Granner D., Andreone, T., Sasaki, K., and Beale, E. (1983). Inhibition of transcription of the phosphoenolpyruvate carboxykinase gene by insulin. *Nature* **305**, 549-551.

Gray, H.B., Ostrander, D.A., Hodnett, J.L., Legerski, R.J., and Robberson, D.L. (1975). Extracellular nucleases of *Pseudomonas* BAL 31. I. Characterization of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities. *Nucl. Acids Res.* **2**, 1459-1492.

Gray, P.W., Leung, D.W., Pennica, D., Yelverton, E., Najarian, R., Simonsen, C.C., Derynck, R., Sherwood, P.J., Wallace, D.M., Berger, S.L., Levinson, A.D., and Goeddel, D.V. (1982). Expression of human immune interferon cDNA in E. coli and monkey cells. *Nature* **295**, 503-508.

Gray, P.W. and Goeddel, D.V. (1983). Cloning and expression of murine immune interferon cDNA. *Proc. Natl. Acad. Sci. USA* **80**, 5842-5846.

Green, W.C. and Leonard, W.J. (1986). The human interleukin-2 receptor. *Ann. Rev. Immunol* . **4**, 69-95.

Guo, L., Yang, R.C.A., and Wu, R. (1983). An improved strategy for rapid direct sequencing of both strands of long DNA molecules cloned in a plasmid. *Nucl. Acids Res.* **11**, 5521-5540.

Hanahan, D. (1985). Techniques for transformation of *E. coli*. In *DNA Cloning: A Practical Approach* , D.M. Glover, ed. (Oxford: IRL Press), p. 119.

Harnish, D.G., Bleackley, R.C., Havele, C., Lobe, C.G., and Paetkau, V. (1986). Translocation of murine interleukin 2 into microsomes during translation in a cell-free system. *Molecular Immunol.* **23**, 201-207.

Hastie, N.D. and Bishop, J.O. (1976). The expression of three abundance classes of messenger RNA in mouse tissues. *Cell* **9**, 761-774.

Hastings, K.E.M. and Emerson, C.P. (1982). cDNA clone analysis of six co-regulated

mRNAs encoding skeletal muscle contractile proteins. *Proc. Natl. Acad. Sci. USA* **79**, 1553-1557.

Hedrick, S.M., Cohen, D.I., Nielsen, E.A., and Davis, M.M. (1984a). Isolation of cDNA encoding T cell-specific membrane-associated proteins. *Nature* **308**, 149-153.

Hedrick, S.M., Nielsen, E.A., Kavaler, J., Cohen, D.I., and Davis, M.M. (1984b). Sequence relationships between putative T-cell receptor polypeptides and immunoglobulins. *Nature* **308**, 153-158.

Henney, C.S., Kuribayashi, K., Kern, D.E., and Gillis, S. (1981). Interleukin-2 augments natural killer cell activity. *Nature* **291**, 335-338.

Hilfiker, M., Moore, R.N., and Farrar, J.J. (1981). Biologic properties of chromatographically separated murine thymoma-derived interleukin 2 and colony stimulating factor. *J. Immunol.* **127**, 1983-1987.

Hirschhorn, R.R., Aller, P., Yuan, Z., Gibson, C.W., and Baserga, R. (1984). Cell-cycle-specific cDNAs from mammalian cells temperature sensitive for growth. *Proc. Natl. Acad. Sci. USA* **81**, 6004-6008.

Hood, L.E., Wilson, J.H., and Wood, W.B. (1975). In *Molecular Biology of Eucaryotic Cells*. (Menlo Park: W.A. Benjamin).

Hooton, J.W.L., Gibbs, C., and Paetkau, V. (1985). Interaction of interleukin 2 with cells: quantitative analysis of effects. *J. Immunol.* **135**, 2464-2473.

Howard, M., and Paul, W.E. (1982). Interleukins for B lymphocytes. *Lymphokine Res.* **1**, 1-4.

Howard, M., Matis, L., Malek, T.R., Shevach, E., Kell, W., Cohen, D., Nakanishi, K., and Paul, W.E. (1983). Interleukin 2 induces antigen-reactive T cell lines to secrete BCGF-1. *J. Exp. Med*. **158**. 2024-2039.

Hoyt, M.A., Knight, D.M., Das, A., Miller, H.I., and Echols, H. (1982). Control of phage lambda development by stability and synthesis of cII protein: role of the viral *cIII* and host *hflA* , *himA* and *himD* genes. *Cell*, **31**, 565-573.

Huynh, T., Young, R.A., and Davis, R.A. (1985). Constructing and screening cDNA libraries in lambda gt 10 and lambda gt 11. In *DNA Cloning, A Practical Approach*, D.M. Glover, ed. (Oxford: IRL Press).

Inaba, K., Granelli-Piperno, A., and Steinman, R.M. (1983). Dendritic cells induce T lymphocytes to release B cell-stimulating factors by an interleukin 2-dependent mechanism. *J. Exp. Med.* **158**, 2040-2057.

Kamalay, J.C. and Goldberg, R.B. (1980). Regulation of structural gene expression in tobacco. *Cell* **19**, 935-946.

Kashima, N., Nishi-Takaoka, C., Fujita, T., Taki, S., Yamada, G., Hamuro, J., and Taniguchi, T. (1985). Unique structure of murine interleukin-2 as deduced from

cloned cDNAs. *Nature* **313**, 402-404.

Kavathas, P., Sukhatme, V.P., Herzenberg, L.A., and Parnes, J.R. (1984). Isolation of the gene encoding the human T-lymphocyte differentiation antigen Leu-2 (T8) by gene transfer and cDNA subtraction. *Proc. Natl. Acad. Sci. USA* **81**, 7688-7692.

Kennedy, N., Knedlitschek, G., Groner, B., Hynes, N.E., Herrlich, P., Michalides, R., and VanOoyen, A.J.J. (1982). Long terminal repeats of endogenous mouse mammary tumor virus contain a long open reading frame which extends into adjacent sequences. *Nature* **295**, 622-624.

Kishimoto, T. (1985). Factors affecting B-cell growth and differentiation. *Ann. Rev. Immunol.* **3**, 133-157.

Kohne, D.E., Levison, S.A., and Byers, M.J. (1977). Room temperature method for increasing the rate of DNA reassociation by many thousandfold: the phenol emulsion reassociation technique. *Biochemistry* **16**, 5329-5341.

Kronenberg, M., Sui, G., Hood, L., and Shastri, N. (1986). The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Ann. Rev. Immunol.* **4**, 529-591.

Kunkel, L.M., Monaco, A.P., Middlesworth, W., Ochs, H.D., and Latt, S.A. (1985). Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. *Proc. Natl. Acad. Sci. USA* **82**, 4778-4782.

Kwon, B.S. and Weissman, S.M. (1984). Mouse mammary tumor virus-related sequences in mouse lymphocytes are inducible by 12-O-tetradecanoyl phorbol-13-acetate. *J. Virol.* **52**, 1000-1004.

Lamar, E.E. and Palmer, E. (1984). Y-encoded, species-specific DNA in mice: evidence that the Y chromosome exists in two polymorphic forms in inbred strains. *Cell* **37**, 171-177.

Langridge, J., Langridge, P., and Bergquist, P.L. (1980). Extraction of nucleic acids from agarose gels. *Analytical Biochem.* **103**, 264-271.

Larsson, E.L. (1980). Cyclosporin A and dexamethasone suppress T cell responses by selectively acting at distinct sites of the triggering process. *J. Immunol.* **124**, 2828-2833.

Lasky, L.A., Lev, Z., Xin, J., Britten, R.J., and Davidson, E.H. (1980). Messenger RNA prevalence in sea urchin embryos measured with cloned cDNAs. *Proc. Natl. Acad. Sci. USA* **77**, 5317-5321.

Lee, F., Yokota, T., Otsuka, T., Gemmell, L., Larson, N., Luh, J., Aria, K., and Rennick, D. (1985). Isolation of cDNA for a human granulocyte-macrophage colony stimulating factor by functional expression in mammalian cells. *Proc. Natl. Acad. Sci. USA* **82**, 4360-4364.

Lee, F., Yokota, T., Otsuka, T., Meyerson, P., Villaret, D., Coffman, R., Mosmann,

T., Rennick, D., Roehm, N., Smith, C., Zlotnik, A., and Arai, K. (1986). Isolation and characterization of a mouse interleukin cDNA clone that expresses B-cell stimulatory factor 1 activities and T-cell and mast-cell-stimulating activities. *Proc. Natl. Acad. Sci. USA* **83**, 2061-2065.

Leonard, W.J., Kronke, M., Peffer, N.J., Depper, J.M., and Greene, W.C. (1985). Interleukin 2 receptor gene expression in normal human T lymphocytes. *Proc. Natl. Acad. Sci. USA* **82**, 6281-6285.

Lewin, B. (1980). In *Gene Expression* Vol. 2., 2nd edn. (New York: Wiley).

Linzer, D.H.I. and Nathans, D. (1983). Growth-related changes in specific mRNAs of cultured mouse cells. *Proc. Natl. Acad. Sci. USA* **80**, 4271-4275.

Littman, D.R., Thomas, Y., Maddon, P.J., Chess, L., and Axel, R. (1985). The isolation and sequence of the gene encoding T8: a molecule defining functional classes of T lymphocytes. *Cell* **40**, 237-246.

Lobe C.G., Havele, C., and Bleackley, R.C. (1986). Cloning of two genes that are specifically expressed in activated cytotoxic T lymphocytes. *Proc. Natl. Acad. Sci. USA* **83**, 1448-1452.

Lopata, M.A., Cleveland, D.W., and Sollner-Webb, B. (1984). High level transient expression of a chloramphenicol acetyl transferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. *Nucl. Acids Res.* **12**, 5707-5717.

Lund, P.K., Moats-Staats, B.M., Simmons, J.G., Hoyt, E., D'Ercole, A.J., Martin, F., and Van Wyk, J.J. (1985). Nucleotide sequence analysis of a cDNA encoding human ubiquitin reveals that ubiquitin is synthesized as a precursor. *J. Biol. Chem.* **260**, 7609-7613.

Maddon, P.J., Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L., and Axel, R. (1985). The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. *Cell* **42**, 93-104.

Mangiarotti, G., Chung, S., Zucker, C., and Lodish, H.F. (1981). Selection and analysis of cloned developmentally-regulated *Dictyostelium discoideum* genes by hybridization-competition. *Nucl. Acids Res.* **9**, 947-963.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). *Molecular Cloning, a Laboratory Manual.* (Cold Spring Harbor: Cold Spring Harbor Laboratories).

Mather, E.L., Alt, F.W., Bothwell, A.L.M., Baltimore, D., and Koshland, M.E. (1981). Expression of J chain RNA in cell lines representing different stages of B lymphocyte differentiation. *Cell* **23**, 369-378.

Messing, J. (1983). New M13 vectors for cloning. *Methods Enzymol.* **101**, 20-79.

Metcalf, D. (1985). The granulocyte-macrophage colony-stimulating factors. *Science* **229**, 16-22.

Metcalf, D. (1986). The molecular biology and functions of the granulocyte-macrophage colony-stimulating factors. *Blood* **67**, 257-267.

Michalides, R., Wagenaar, E., and Weijers, P. (1985). Rearrangements in the long terminal repeat of extra mouse mammary tumor proviruses in T-cell leukemias of mouse strain GR result in a novel enhancer-like structure. *Molec. Cell Biol.* **5**, 823-830.

Milanese, C., Siliciano, R.F., Schmidt, R.E., Ritz, J., Richardson, N.E., and Reinherz, E.L. (1986). A lymphokine that activates the cytotolytic program of both cytotoxic T lymphocytes and natural killer clones. *J. Exp. Med.* **163**, 1583-1588.

Miller, J., Malek, T.R., Leonard, W.J., Greene, W.C., Shevach, E.M., and Germain, R.N. (1985). Nucleotide sequence and expression of a mouse interleukin 2 receptor cDNA. *J. Immunol.* **134**, 4212-4217.

Milner, R.J. and Sutcliffe, J.G. (1983). Gene expression in rat brain. *Nucl. Acids Res.* **11**, 5497-5520.

Miyatake, S., Otsuka, T., Yokota, T., Lee, F., and Arai, K. (1985). Structure of the chromosomal gene for granulocyte-macrophage colony stimulating factor: comparison of the mouse and human genes. *EMBO J.* **4**, 2562-2568.

Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* **65**, 55-63.

Mosmann, T., Cherwinski, H., Cher, D., Bond, M., and Coffman, R. (1986). Two types of mouse helper T cell clones differ in function and produce distinct sets of lymphokines. Abstract 3.21.23, Sixth International Congress of Immunology, p. 299.

Moss, B. and Koczot, F. (1976). Sequence of methylated nucleotides at the 5' terminus of adenovirus-specific RNA. *J. Virol.* **17**, 385-392.

Nakamura, M., Manser, T., Pearson, G.D.N., Daley, J.J., and Gefter, M.L. (1984). Effects of INF-gamma on the immune response *in vivo* and on gene expression *in vitro*. Nature **307**, 381-382.

Ng, J. (1984) M. Sc. Thesis. (Edmonton: University of Alberta).

Nishizuka, Y. (1984). The role of protein kinase C in cell surface signal transduction and tumor promotion. *Nature* **308**, 693-698.

Noma, Y., Sideras, P., Naito, T., Bergstedt-Lindquist, S., Azuma, C., Severinson, E., Tanabe, T., Kinashi, T., Matsuda, F., Yaoita, Y., and Honjo, T. (1986). Cloning of cDNA encoding the murine IgG1 induction factor by a novel strategy using SP6 promoter. *Nature* **319**, 640-646.

Orosz, C.G., Fidelus, R.K., Roopenian, D.C., Widmer, M.B., Ferguson, R.M., and

Bach, F.H. (1982). Analysis of cloned T cell function. 1. Dissection of cloned T cell proliferative responses using cyclosporin A. *J. Immunol.* **129**, 1865-1868.

Paetkau, V.P., Bleackley, R.C., Riendeau, D., Harnish, D.G., and Holowachuk, E.W. (1984). Toward the molecular biology of IL2. In *Contemporary Topics in Molecular Immunology* , **10**, S. Gillis and F.P. Inman, eds. (New York: Plenum Press), pp. 35-61.

Paetkau, V., Shaw, J., Ng, J., Elliott, J.F., Meerovitch, K., Barr, P.J., and Bleackley, R.C. (1985). Regulation of lymphokine expression. In: *Mediators of Immune Regulation and Immunotherapy*. K. Singhal and N.R. Sinclair, eds. (New York: N.R. Elsevier Press).

Paul, W.E. (1984). The immune system: an introduction. In *Fundamental Immunology*, W.E. Paul, ed. (New York: Raven Press).

Paulnock-King, D., Sizer, K.C., Freund, Y.R., and Jones, P.P. (1985). Coordinate induction of Ia alpha, beta, and Ii mRNA in a macrophage cell line. *J. Immunol.* **135**, 632-636.

Peterson, D.O., Kriz, K.G., Marich, J.E., and Toohey, M.G. (1985). Sequence organization and molecular cloning of mouse mammary tumor virus DNA endogenous to C57Bl/6 mice. *J. Virol.* **54**, 525-531.

Racevskis, J. (1986). Expression of the protein product of the mouse mammary tumor virus long terminal repeat gene in phorbol ester-treated mouse T-cell-leukemia cells. *J. Virol.* **58**, 441-449.

Ramirez, F., Natta, C., O'Donnell, J.V., Canale, V., Bailey, G., Sanguensermsri, T., Maniatis, G.M., Marks, P.A., and Bank, A. (1975). Relative numbers of human globin genes assayed with purified alpha and beta complementary human DNA. *Proc. Nat. Acad. Sci. USA* **72**, 1550-1554.

Redmond, S.M.S., and Dickson, C. (1983). Sequence and expression of the mouse mammary tumor virus env gene. *EMBO J.* **2**, 125-131.

Reem, G.H., Cook, L.A., and Vilcek, J. (1983). Gamma interferon synthesis by human thymocytes and T lymphocytes inhibited by cyclosporin A. *Science* **221**, 63-65.

Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**, 237-251.

Rosenberg, S.A., Spiess, P., and Lafreniere, R. (1986). A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes. *Science* **233**, 1318-1321.

Rowekamp, W. and Firtel, R.A. (1980). Isolation of developmentally regulated genes from *Dictyostelium* .. *Develop. Biol.* **79**, 409-418.

Saito, H., Kranz, D.M., Takagaki, Y., Hayday, A.C., Eisen, H.N., and Tonegawa,

S. (1984a). Complete primary structure of a heterodimeric T-cell receptor deduced from cDNA sequences. *Nature* **309**, 757-762.

Saito, H., Kranz, D.M., Takagaki, Y., Hayday, A.C., Eisen, H.N., and Tonegawa, S. (1984b). A third rearranged and expressed gene in a clone of cytotoxic T lymphocytes. *Nature* **312**, 36-40.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Sargent, T.D. and Dawid, I.B. (1983). Differential gene expression in the gastrula of *Xenopus laevis* . *Science* **222**,135-139.

Schutzbank, T., Robinson, R., Oren, M., and Levine, A.J. (1982). SV40 large tumor antigen can regulate some cellular transcripts in a positive fashion. *Cell* **30**, 481-490.

Schwartz, R.H. (1984). The role of gene products of the major histocompatibility complex in T cell activation and cellular interactions. In *Fundamental Immunology*, W.E. Paul, ed. (New York: Raven Press), pp. 379-438.

Seligy, V.L., Barbier, J.R., Dimock, K.D., Dove, M.J., Moranelli, F., Morosoli, R., Willick, G.E., and Yaguchi, M. (1983). Current status in the application of recombinant DNA technology in the construction of cellulolytic yeast strains. *Foundations for Biotechnical and Industrial Fermentation Research* **1**, 167-185.

St. John, T.P. and Davis, R.W. (1979). Isolation of galactose-inducible DNA sequences from *Saccharomyces cerevisiae* by differential plaque filter hybridization. *Cell* **16**, 443-452.

Stanley, E., Metcalf, D., Sobieszczuk, P., Gough, N.M., and Dunn, A.R. (1985). The structure and expression of the murine gene encoding granulocyte-macrophage colony stimulating factor: evidence for utilisation of alternative promoters. *EMBO J.* **10**, 2569-2573.

Steeg, P.S., Moore, R.N., Johnson, H.M., and Oppenheim, J.J. (1982). Regulation of murine macrophage Ia antigen expression by a lymphokine and immune interferon activity. *J. Exp. Med.* **156**, 1780-1793.

Stehelin, D., Guntaka, R.V., Varmus, H.E., and Bishop, J.M. (1976). Purification of DNA complementary to nucleotide sequences required for neoplastic transformation of fibroblasts by avian sarcoma viruses. *J. Mol. Biol.* **101**, 349-365.

Stetler, G.L. and Thorner, J. (1984). Molecular cloning of hormone-responsive genes from the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **81**, 1144-1148.

Tada, T. (1984).Help, suppression, and specific factors. In *Fundamental Immunology*, W.E. Paul, ed. (New York: Raven Press), pp. 481-518.

Taniguchi, T., Sakai, M., Fujii-Kuriyama, Y., Muramatsu, M., Kobayashi, S., and

Sudo, T. (1979). Construction and identification of a bacterial plasmid containing the human fibroblast interferon gene sequence. *Proc. Jpn. Acad.* **55**, Ser. B. 464-469.

Taniguchi, T., Fujii-Kuriyama, Y., and Muramatsu, M. (1980). Molecular cloning of human interferon cDNA. *Proc. Natl. Acad. Sci. USA* **77**, 4003-4006.

Taniguchi, T., Matsui, H., Fujita, T., Takoka, C., Kashima, N., Yoshimoto, R., and Hamuro, J. (1983). Structure and Expression of a cloned cDNA for human interleukin-2. *Nature* **302**, 305-310.

Tesch, H., Muller, W., and Rajewsky, K. (1986). Lymphokines regulate immunoglobulin isotype expression in an antigen-specific immune response. *J. Immunol.* **136**, 2892-2895.

Thomas, P.S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* **77**, 5201-5205.

Timberlake, W.E. (1980). Developmental gene regulation in *Aspergillus nidulans*. *Develop. Biol.* **78**, 497-510.

Trueh, A., Albert, F., Goldstein, P., and Schmitt-Verhulst, A. (1985). Early steps of lymphocyte activation bypassed by synergy between calcium ionophores and phorbol ester. *Nature* **313**, 318-320.

Van den Elsen, P., Shepley, B., Borst, J., Coligan, J.E., Markham, A.F., Orkin, S., and Terhorst, C. (1984). Isolation of cDNA clones encoding the 20K T3 glycoprotein of human T-cell receptor complex. *Nature* **312**, 413-418.

Van Ooyen, A.J.J., Michalides, R.J.A.M., and Nusse, R. (1983). Structural analysis of a 1.7-kilobase mouse mammary tumor virus-specific RNA. *J.Virol.* **46**, 362-370.

Wagner, H. and Rollinghoff, M. (1978). T-T cell interactions during in vitro cytotoxic allograft responses. I. Soluble product from activated Ly3+ T cells trigger autonomously antigen-primed Ly12+ T cells to proliferate and show cytotoxic activity. *J. Exp. Med.* **148**, 1523-1538.

Weinberger, J., Baltimore D., and Sharp, P.A. (1986). Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. *Nature* **322**, 846-848.

Weiss, A., Wiskocil, R.L., and Stobo, J.D. (1984a). The role of T3 surface molecules in the activation of human T cells: a two stimulus requirement for IL-2 production reflects events occurring at a pretranslational level. *J. Immunol.* **133**, 123-128.

Weiss, A., Imboden, J., Shoback, D., and Stobo, J. (1984b). Role of T3 surface molecules in human T cell activation: T3-dependent activation results in an increase in cytoplasmic-free calcium. *Proc. Natl. Acad. Sci. USA* **81**, 4169-4173.

Weiss, A., Imboden, J., Hardy, K., Manger, B., Terhorst, C., and Stobo, J. (1986).

The role of T3/antigen receptor complex in T-cell activation. *Ann. Rev. Immunol.* **4**, 593-619.

Wellinger, R.J., Garcia, M., Vessaz, A., and Diggelmann, H. (1986). Exogenous mouse mammary tumor virus proviral DNA isolated from a kidney adenocarcinoma cell line contains alterations in the U3 region of the long terminal repeat. *J. Virol.* **60**, 1-11.

Wheeler, D.A., Butel, J.S., Medina, D., Cardiff, R.D., and Hager, G.L. (1983). Transcription of mouse mammary tumor virus: identification of a candidate mRNA for the long terminal repeat gene product. *J. Virol.* **46**, 42-49.

White, B.A. and Bancroft, F.C. (1982). Cytoplasmic dot hybridization. *J. Biol. Chem.* **257**, 8569-8572.

Williams J.G. and Lloyd, M.M. (1979). Changes in the abundance of polyadenylated RNA during slime mould development measured using cloned molecular hybridization probes. *J. Mol. Biol.* **129**, 19-35.

Wittek, R., Cooper, J.A., Barbosa, E., and Moss, B. (1980). Expression of the vaccinia virus genome: analysis and mapping of mRNAs encoded within the inverted terminal repetition. *Cell* **21**, 487-493.

Wong, G.G., Witek, J., Temple, P.A., Wilkens, K.M., Leary, A.C., Luxenberg, D.P., Jones, S.S., Brown, E.C., Kay, R.M., Orr, E.C., Shoemaker, C., Golde, D.N., Kaufman, R.J., Hewick, R.M., Wang, E.A., and Clark, S.C. (1985). Human GM-CSF: molecular cloning of the cDNA and purification of the natural and recombinant proteins. *Science* **228**, 810-815.

Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S.P., Aleksander, I., and Mak, T.W. (1984). A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. *Nature* **308**, 145-149.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

Yokota, T., Lee, F., Rennick, D., Hall, C., Arai, N., Mosmann, T., Nabel, G., Cantor, H., and Arai, K. (1984). Isolation and characterization of a mouse cDNA clone that expresses mast-cell growth-factor activity in monkey cells. *Proc. Natl. Acad. Sci. USA* **81**, 1070-1074.

Young, J.D. and Cohn, Z.A. (1986). Cell-mediated killing: a common mechanism? *Cell* **46**, 641-642.

Zimmermann, C.R., Orr, W.C., Leclerc, R.F., Barnard, E.C., and Timberlake, W.E. (1980). Molecular cloning and selection of genes regulated in *Aspergillus* development. *Cell* **21**, 709-715.